

97526

STIC-Biotech/ChemLib

---

**From:** Rao, Manjunath N.  
**Sent:** Thursday, June 26, 2003 11:20 AM  
**To:** STIC-Biotech/ChemLib  
**Subject:** Sequence search request for 09/762,767

**Fr m:** Manjunath N. Rao  
Art Unit 1652, Room 10A11  
Mail Box in Room 10D 01  
Phone: 306-5681

**Date:** 6-26-03

Please search the following as soon as possible for application with serial number **09/762,767**

1. **SEQ ID NO:1 and 3**, against all commercial nucleic acid databases including issued patents database and pending application database and provide a print of all results.
2. **SEQ ID NO: 2 and 4** , against all commercial protein databases including issued patents database and pending application database and provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.  
Biotechnology Patent Examiner  
Art Unit 1652, Room 10A11  
Mail Box in 10D01  
Crystal Mall 1, USPTO.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 27, 2003, 12:42:07 ; Search time 36.3713 Seconds  
(without alignments)  
2527.899 Million cell updates/sec

Title: US-09-762-767A-2  
 Perfect score: 3593  
 Sequence: 1 MCTRLKLDKLTGECPSLL.....QKKDVEDGNANFLGRASGID 690

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

Database : A.GeneSeq.101002.\*

1:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3593	100.0	690	21	AAV70475	Human cyclic nucle
2	3593	100.0	690	21	AAV51607	Human soluble guan
3	3593	100.0	690	23	ABG61940	Prostate cancer-as
4	3171.5	88.3	717	23	ABG51878	Prostate cancer-as
5	991.5	27.6	676	22	ABB58042	Drosophila melanog
6	789.5	22.0	619	21	AAV51608	Human soluble guan
7	743	20.7	787	22	ABB58149	Drosophila melanog
8	642.5	17.9	690	22	ABB60360	Drosophila melanog
9	462	12.9	1076	22	ABB59684	Drosophila melanog
10	461	12.8	1115	22	ABB83940	Amino acid sequenc

11	460	12.8	1025	12	AAR38683	GC- $\beta$ . Rattus ratt
12	460	12.8	1047	14	AAR10369	Human. Natriuretic
13	460	12.8	1047	12	AAR10867	NPR3(Pro655, Glu65
14	457.5	12.7	1029	13	AAR38862	GC- $\beta$ . Rattus ratt
15	456.5	12.7	1061	14	AAU11280	Human atrionatriur
16	456.5	12.7	1061	23	AAU11281	Human atrionatriur
17	456.5	12.7	1075	22	ABU11783	Human ANP-A-recept
18	456.5	12.7	1172	22	ABH72034	Drosophila melanog
19	453	12.6	1076	22	ABH71206	Drosophila melanog
20	445	12.4	975	22	ABH64661	Drosophila melanog
21	423.5	11.8	1525	22	ABH71674	Drosophila melanog
22	421.5	11.7	566	22	ABH61743	Drosophila melanog
23	408.5	11.4	887	22	ABG03067	Novel human diagno
24	408.5	11.4	1073	18	AAW32063	Human $\beta$ 1 receptor
25	408.5	11.4	1073	19	AAW37371	Human $\beta$ 2 receptor
26	408.5	11.4	1073	23	AAU08788	Human $\beta$ 3 receptor
27	407.5	11.3	1075	14	AAR38681	Human guanylin cyc
28	403.5	11.2	1151	22	AAH83939	GC- $\beta$ . Rattus ratt
29	398.5	11.1	949	22	AAH83941	Amino acid sequenc
30	397	11.0	669	22	ABH65903	Amino acid sequenc
31	396	11.0	649	22	ABH65902	Drosophila melanog
32	356	9.9	273	22	AAH67400	Drosophila melanog
33	325.5	9.1	1077	22	AAU01924	Partial amino acid
34	322.5	9.0	887	23	AAE14741	Human adenylate cy
35	322.5	9.0	1086	22	AAE02938	Human adenylate cy
36	322.5	9.0	1086	23	AAE14742	Human adenylate cy
37	322	9.0	1064	21	AAH02008	Human adenylate cy
38	319.5	8.9	1099	21	AAH02011	Type IV adenylatyl c
39	319	8.9	1090	17	AAR94560	Type VII adenylatyl c
40	319	8.9	1090	17	AAH02006	Rat adenylatyl cycla
41	312	8.7	1086	23	AAH17129	Adenylatyl cyclase t
42	310.5	8.6	1103	23	ABH81465	Human adenylatyl and
43	303.5	8.4	855	21	AAH54359	Human adenylate cy
44	303.5	8.4	857	23	ABG64609	Human pancreatic c
45	303.5	8.4	857	23	ABG64609	Human albumin fusi
					Human gene 11 enco	

## ALIGNMENTS

```

RESULT 1
AAV70475
ID   AAV70475 standard; Protein; 690 AA

```

DT 04-JUL-2000 (first entry)

Human cyclic nucleotide-associated protein-3 (CNAP-3)

KM Cyclic nucleotide-associated protein-3; cNAP-3 human; cycliclastic;  
 KM anti-arteriosclerotic; hepatotropic; anti-leukaemic; anti-inflammatory;  
 KM immunomodulatory; anti-asthmatic; anti-anemic; anti-diabetic; diagnosis;  
 KM antihemolytic; dematologic; neuroprotective; anti-epileptic; cancer;  
 KM anti-Alzheimer's; anti-Parkinsonian; cerebroprotective; ophthalmological;  
 KM anti-fertility; anti-allergic; vasotropic; immunosuppressive;  
 KM hypotensive; gene therapy; prevention; treatment; arteriosclerosis;  
 KM cell proliferative disorder; autoimmune/inflammatory; diabetes mellitus;  
 KM neurological; vision; reproductive; smooth muscle.

**Homo sapiens.**

FT	Key	Location/Qualifiers
FT	Modified-site	65 "/note="Potential phosphorylation site"
FT	Modified-site	211 "/note="Potential phosphorylation site"
FT	Modified-site	361 "/note="Potential phosphorylation site"
FT	Modified-site	450 "/note="Potential phosphorylation site"
FT	Modified-site	478 "/note="Potential phosphorylation site"

ET

FT Modified-site 29 /note= "Potential phosphorylation site"  
 FT Modified-site 73 /note= "Potential phosphorylation site"  
 FT Modified-site 117 /note= "Potential phosphorylation site"  
 FT Modified-site 135 /note= "Potential phosphorylation site"  
 FT Modified-site 158 /note= "Potential phosphorylation site"  
 FT Modified-site 229 /note= "Potential phosphorylation site"  
 FT Modified-site 511 /note= "Potential phosphorylation site"  
 FT Modified-site 527 /note= "Potential phosphorylation site"  
 FT Modified-site 638 /note= "Potential phosphorylation site"  
 FT Modified-site 687 /note= "Potential phosphorylation site"  
 FT Modified-site 39 /note= "Potential phosphorylation site"  
 FT Modified-site 64 /note= "Potential phosphorylation site"  
 FT Modified-site 105 /note= "Potential phosphorylation site"  
 FT Modified-site 261 /note= "Potential phosphorylation site"  
 FT Modified-site 328 /note= "Potential phosphorylation site"  
 FT Modified-site 356 /note= "Potential phosphorylation site"  
 FT Modified-site 360 /note= "Potential phosphorylation site"  
 FT Modified-site 394 /note= "Potential phosphorylation site"  
 FT Modified-site 622 /note= "Potential phosphorylation site"  
 FT Modified-site 635 /note= "Potential phosphorylation site"  
 FT Modified-site 27 /note= "Potential phosphorylation site"  
 FT Modified-site 332 /note= "N-glycosylated"  
 FT Modified-site 479 /note= "N-glycosylated"  
 FT Modified-site 600 /note= "N-glycosylated"  
 FT Modified-site 472..661 /note= "N-glycosylated"  
 FT Region /note= "Guanylate cyclase signature sequence identified by PFAM analysis; CNAP-3 shares 89% identity with human, soluble guanylate cyclase large subunit"  
 XX WO200014248-A1.  
 PD 16-MAR-2000.  
 PF 03-SEP-1999; 99WO-US20287.  
 XX 04-SEP-1998; 98US-0148904.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Hillman JL, Yue H, Guegler KJ, Corley NC, Patterson C, Tang YF.  
 DR WPI: 2000-256994/22.  
 DR N-PSDB: AAZ51684.  
 XX Isolated cyclic nucleotide associated proteins useful for preventing, diagnosing and treating cell proliferative, autoimmune/inflammatory, neurological, vision, reproductive and smooth muscle disorders -  
 XX

PS Disclosure: Page 67-69; 78pp; English.  
 XX  
 CC The present sequence is a human cyclic nucleotide  
 CC associated protein-3 (CNAP-3), identified in incyte clone 159278,  
 CC that is isolated from ADENINB01 cDNA library. It is expressed in  
 CC nervous, reproductive, cardiovascular and developmental tissues.  
 CC CNAP sequences may be used for prevention, treatment and diagnosis of  
 CC diseases associated with altered CNAP expression such as, cell  
 CC proliferative disorders (e.g. arteriosclerosis, cirrhosis, leukemia,  
 CC lymphoma and cancer of the breast, prostate, lung and brain), autoimmune/  
 CC inflammatory disorders (e.g. asthma, anemia, diabetes mellitus, multiple  
 CC sclerosis and psoriasis), neurological disorders (e.g. epilepsy,  
 CC Alzheimer's/Parinson's disease and strokes), vision disorders (e.g.  
 CC conjunctivitis, glaucoma, cataracts and retinitis pigmentosa),  
 CC reproductive disorders (e.g. infertility, uterine fibroids, ectopic  
 CC pregnancies and impotence) and smooth muscle disorders (e.g. angina,  
 CC anaphylactic shock, Keans-Sayre syndrome and hypertension). The  
 CC coding sequence can be used for gene therapy.  
 XX  
 SQ Sequence 690 AA:  
 Query Match 100.0%; Score 3593; DB 21; Length 690;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MCTRLKDKLTGCEPFLAPGVNSESSEBAAGSSCKATVPICODIEKNIQESLP 60  
 1 MCTRLKDKLTGCEPFLAPGVNSESSEBAAGSSCKATVPICODIEKNIQESLP 60  
 61 QKTSRSRYVLTAESECKLFFPERLNVALQTLAKHKIKESKLEDEPKTAE 120  
 61 QKTSRSRYVLTAESECKLFFPERLNVALQTLAKHKIKESKLEDEPKTAE 120  
 121 QAVAGVVEVYKESIGEVEFKICEEDENILGVGTLKDFLNFSTILKQSSHCQAG 180  
 121 QAVAGVVEVYKESIGEVEFKICEEDENILGVGTLKDFLNFSTILKQSSHCQAG 180  
 181 KRGLEDASILCLDEKEDFLHYVYFPKRTSLIPGIIKAAHYLYTEVEVSLMPCF 240  
 181 KRGLEDASILCLDEKEDFLHYVYFPKRTSLIPGIIKAAHYLYTEVEVSLMPCF 240  
 241 HNDSEFYNOPILLYSYMKSTKPSLSPKQSSSVITFSECTPPHFMEDMTILQ 300  
 241 HNDSEFYNOPILLYSYMKSTKPSLSPKQSSSVITFSECTPPHFMEDMTILQ 300  
 301 FNGGIRRLANRRDFOGKNPEEYFEILTPKINQPSGIMTMNQVYVRRMDSYKKS 360  
 301 FNGGIRRLANRRDFOGKNPEEYFEILTPKINQPSGIMTMNQVYVRRMDSYKKS 360  
 361 SRVMDLKQMIYVSSAILFLGSPCVDRLEDFGTGRLYLSDIPHNALRDVYLIGEQR 420  
 361 SRVMDLKQMIYVSSAILFLGSPCVDRLEDFGTGRLYLSDIPHNALRDVYLIGEQR 420  
 421 AODGLKRLGKTLKATLEAQAALBEKKKTYDLCSITPCVAAQOLMGQVYVAAKPSNV 480  
 421 AODGLKRLGKTLKATLEAQAALBEKKKTYDLCSITPCVAAQOLMGQVYVAAKPSNV 480  
 481 TMLFSDIYGTALCSQSPLOVITMTALYFRPDQCGELDYVKETIGDAYCYAGGLHK 540  
 481 TMLFSDIYGTALCSQSPLOVITMTALYFRPDQCGELDYVKETIGDAYCYAGGLHK 540  
 541 ESDTHAVQIALMALKMELSDSVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 541 ESDTHAVQIALMALKMELSDSVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 601 VTANKFESGCVPRKINVSPTTYRLKDCPQVFTPRREELPRPSEIIGICHFLDAY 660  
 601 VTANKFESGCVPRKINVSPTTYRLKDCPQVFTPRREELPRPSEIIGICHFLDAY 660  
 661 OOGTNSKPCFOKKNVEDGNANFLKASGID 690  
 661 OOGTNSKPCFOKKNVEDGNANFLKASGID 690



RESULT 2  
AAVS1607  
ID AAVS1607 standard; Protein: 690 AA.  
XX  
XX AAVS1607;  
AC  
XX 26-MAY-2000 (first entry)  
DT  
XX Human soluble guanylyl cyclase alpha protein.  
XX  
XX  
XX Human: guanylyl cyclase alpha; hsgcalpha; hsgcbeta; soluble;  
KM guanylyl cyclase beta; antiatherosclerotic; vasotropic; hypotensive;  
KM gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;  
KW peripheral arterial occlusive disease; arterial hypertension.  
XX  
XX Homo sapiens.  
OS  
XX DEL9837015-A1.  
XX  
XX 24-FEB-2000.  
PD  
XX  
XX 14-AUG-1998; 98DE-1037015.  
PF  
XX  
XX 14-AUG-1998; 98DE-1037015.  
PR  
XX  
XX (VASO-) VASOPHARM BIOTECH GMBH & CO KG.  
PA  
XX  
XX Schmidt H, Zabel U, Poller W;  
PI  
XX WPI, 2000-184044/17.  
DR  
XX N-PSDB; AA288938.  
XX  
XX  
XX New human soluble guanylate cyclase alpha/beta and the nucleic acid  
PT encoding the subunits, useful for producing diagnostic antibodies, and  
PT for somatic gene therapy of arteriosclerosis -  
XX  
XX Example 1; Page 13-15; 44pp; German.  
XX  
XX This invention describes novel purified human soluble guanylate cyclase  
CC alpha/beta (hsgcalpha/beta). The products of the invention have  
CC antiatherosclerotic, vasotropic and hypotensive activity. Nucleic acid  
CC sequences encoding the alpha and/or beta subunit are useful for somatic  
CC gene therapy of arteriosclerosis and restenosis, ischemia (infarct),  
CC peripheral arterial occlusive disease and arterial hypertension.  
CC Antibodies to hsgcalpha/beta can be used for diagnosis of aberrant  
CC hsgcalpha/beta expression in human tissues. This sequence represents  
CC the human soluble guanylyl cyclase alpha subunit described in the method  
CC of the invention.  
XX  
XX Sequence 690 AA:  
SQ  
Query Match 100.0%; Score 3593; DB 21; Length 690;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 HNDCESEVNPYLLYSVHMASTKPSLSPKQSSLVPTSLFCKTFPHFMDKMTIILQ 300  
QY 301 FGNIRRLANRRPOGKPNFEFEITLPKINOTFSGIMTMANOFVVRBRDMSYKKS 360  
DB 301 FGNIRRLANRRPOGKPNFEFEITLPKINOTFSGIMTMANOFVVRBRDMSYKKS 360  
QY 361 SRVMDLKGOMIYVESAILFLGSPCYDRLEDFGTGRGLYLSIDIPINHALDVLIGBOAR 420  
DB 361 SRVMDLKGOMIYVESAILFLGSPCYDRLEDFGTGRGLYLSIDIPINHALDVLIGBOAR 420  
QY 421 AOGGLKRLKLTATLQAOALFEKKKTVDLCSIFPCVVAQOLMGQVQAKKFSNV 480  
DB 421 AOGGLKRLKLTATLQAOALFEKKKTVDLCSIFPCVVAQOLMGQVQAKKFSNV 480  
QY 481 TMLFSDIVGFTALCSQSPLOVITMLNALTTRPDQCGELDYKQVETIGAYCAGGLHK 540  
DB 481 TMLFSDIVGFTALCSQSPLOVITMLNALTTRPDQCGELDYKQVETIGAYCAGGLHK 540  
QY 541 ESDTHAVQIALMALKMEISDEVMSPHGEPIKMRIGLHSGSVAGVYVMPRYCLFGNN 600  
DB 541 ESDTHAVQIALMALKMEISDEVMSPHGEPIKMRIGLHSGSVAGVYVMPRYCLFGNN 600  
QY 601 VTLANKFESCSVPKINVSPTTYRLKDCQGFYFTPPSRBELPPNPSLPGICHTDAY 660  
DB 601 VTLANKFESCSVPKINVSPTTYRLKDCQGFYFTPPSRBELPPNPSLPGICHTDAY 660  
QY 661 OQGTNSKRCFOKQKDVEDGNANFLGKASGID 690  
DB 661 OQGTNSKRCFOKQKDVEDGNANFLGKASGID 690

RESULT 3  
ABG61940  
ID ABG61940 standard; Protein: 690 AA.  
XX  
XX ABG61940;  
AC  
XX 15-AUG-2002 (first entry)  
DT  
XX  
XX Prostate cancer-associated protein #141.  
DE  
XX  
XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.  
KM  
XX  
XX Mammalia.  
OS  
XX  
XX W02002030268-A2.  
PN  
XX  
XX 18-APR-2002.  
PD  
XX  
XX 12-OCT-2001; 2001WO-US32045.  
PF  
XX  
XX 13-OCT-2000; 2000US-0687576.  
PR 08-DEC-2000; 2000US-0733288.  
PR 08-DEC-2000; 2000US-0733288.  
PR 24-JAN-2001; 2001US-2639572.  
PR 16-MAR-2001; 2001US-276791P.  
PR 16-MAR-2001; 2001US-276888P.  
PR 06-APR-2001; 2001US-281922P.  
PR 24-APR-2001; 2001US-286214P.  
PR 30-APR-2001; 2001US-0847046.  
PR 04-MAY-2001; 2001US-288589P.  
XX  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
PA  
XX  
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevez P;  
FI  
XX WPI; 2002-471335/50.  
DR  
XX N-PSDB; ABK92259.  
DR  
XX  
XX Detecting a prostate cancer-associated transcript in a cell in a  
PT patient, useful for diagnosing prostate cancer (PC) or screening  
PT modulators of PC, by determining if prostate cancer-associated genes

```

PT are expressed in a prostate tissue
XX
XX Claim 27; Page 421-422; 436pp; English.
CC The present invention relates to methods of detecting a prostate
CC cancer-associated transcript in a cell from a patient. The method
CC comprises contacting a biological sample from the patient with
CC prostate cancer-associated polynucleotides (designated PC genes) that
CC selectively hybridise to a sequence that is at least 80% identical
CC to them. The prostate cancer-associated polynucleotide sequences
CC are differentially expressed in prostate tumour tissue or in
CC prostate cancer and are derived from the tissues of various
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC The methods of the invention are useful for diagnosing and treating
CC prostate cancer in mammals. The prostate cancer-associated genes are
CC useful for diagnosing or treating prostate cancer, as well as for
CC identifying modulators of prostate cancer or agents that inhibit
CC prostate cancer. The nucleic acid sequences are particularly useful
CC in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins.
XX
XX Sequence 690 AA:
SQ
Query Match 100.0%; Score 3593; DB 23; Length 690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MECTRLDKLKITGECFSLAPGVNPNSSSEAGSSSSCKATVPICODIPKNIQESLP 60
DB 1 MECTRLDKLKITGECFSLAPGVNPNSSSEAGSSSSCKATVPICODIPKNIQESLP 60
OY 61 ORKTSRSVYVHTLAESTCKLFFPEERLNAVALQTLAKHKIKESKSLERDEPKTAE 120
DB 61 ORKTSRSVYVHTLAESTCKLFFPEERLNAVALQTLAKHKIKESKSLERDEPKTAE 120
OY 121 QAVAGVVEVYKESIEEVEFKICEEDENILGVVGTGLKDFLNFSTLLKSSCOENG 180
DB 121 QAVAGVVEVYKESIEEVEFKICEEDENILGVVGTGLKDFLNFSTLLKSSCOENG 180
OY 121 QAVAGVVEVYKESIEEVEFKICEEDENILGVVGTGLKDFLNFSTLLKSSCOENG 180
DB 121 QAVAGVVEVYKESIEEVEFKICEEDENILGVVGTGLKDFLNFSTLLKSSCOENG 180
OY 181 KRGRLDASILCLDKEEDFLHAYVFPKRTTSLIPGIIKAAHVLVTEVEVSLMPCE 240
DB 181 KRGRLDASILCLDKEEDFLHAYVFPKRTTSLIPGIIKAAHVLVTEVEVSLMPCE 240
OY 181 KRGRLDASILCLDKEEDFLHAYVFPKRTTSLIPGIIKAAHVLVTEVEVSLMPCE 240
DB 181 KRGRLDASILCLDKEEDFLHAYVFPKRTTSLIPGIIKAAHVLVTEVEVSLMPCE 240
OY 241 HNDSEFVNOPYLLSYVHMKSTKPSLSPKQSSIVTPTSLFCKFPFHMDKMTLIQ 300
DB 241 HNDSEFVNOPYLLSYVHMKSTKPSLSPKQSSIVTPTSLFCKFPFHMDKMTLIQ 300
OY 241 HNDSEFVNOPYLLSYVHMKSTKPSLSPKQSSIVTPTSLFCKFPFHMDKMTLIQ 300
DB 241 HNDSEFVNOPYLLSYVHMKSTKPSLSPKQSSIVTPTSLFCKFPFHMDKMTLIQ 300
OY 301 FGNGIRRLANRRDFQGNPFEEFEILPKINQTFSGIMTMNQFVVRVRMDSVKRS 360
DB 301 FGNGIRRLANRRDFQGNPFEEFEILPKINQTFSGIMTMNQFVVRVRMDSVKRS 360
OY 301 FGNGIRRLANRRDFQGNPFEEFEILPKINQTFSGIMTMNQFVVRVRMDSVKRS 360
DB 301 FGNGIRRLANRRDFQGNPFEEFEILPKINQTFSGIMTMNQFVVRVRMDSVKRS 360
OY 361 SRVMDLKQMTIYVSSAILFLGSPCVRLDEDFTRGLYLSDIPIHMLRDVVLIGEOR 420
DB 361 SRVMDLKQMTIYVSSAILFLGSPCVRLDEDFTRGLYLSDIPIHMLRDVVLIGEOR 420
OY 421 AODGKRLGLKATLEQAHQALEEKKKYVDLCSITPCERAOQDMGOVVAQAKFSNV 480
DB 421 AODGKRLGLKATLEQAHQALEEKKKYVDLCSITPCERAOQDMGOVVAQAKFSNV 480
OY 481 TMLPSDIYGFPAISQCSPLQVITMLNLYTRFDQCGELVYKVTIGDAYCVAGGJHK 540
DB 481 TMLPSDIYGFPAISQCSPLQVITMLNLYTRFDQCGELVYKVTIGDAYCVAGGJHK 540
OY 481 TMLPSDIYGFPAISQCSPLQVITMLNLYTRFDQCGELVYKVTIGDAYCVAGGJHK 540
DB 481 TMLPSDIYGFPAISQCSPLQVITMLNLYTRFDQCGELVYKVTIGDAYCVAGGJHK 540
OY 541 ESDTHAVOIALMALKMLSDSVMSPHGEPIKMRIGLHSGSVFAGVGVKMRXYCLFGNN 600
DB 541 ESDTHAVOIALMALKMLSDSVMSPHGEPIKMRIGLHSGSVFAGVGVKMRXYCLFGNN 600
OY 541 ESDTHAVOIALMALKMLSDSVMSPHGEPIKMRIGLHSGSVFAGVGVKMRXYCLFGNN 600
DB 541 ESDTHAVOIALMALKMLSDSVMSPHGEPIKMRIGLHSGSVFAGVGVKMRXYCLFGNN 600
OY 601 VTLANKFSSCVPRKINVSPTTYRLKDCRGVFPFRBRRELPRPFSEIIGICFLDAY 660
DB 601 VTLANKFSSCVPRKINVSPTTYRLKDCRGVFPFRBRRELPRPFSEIIGICFLDAY 660
OY 601 VTLANKFSSCVPRKINVSPTTYRLKDCRGVFPFRBRRELPRPFSEIIGICFLDAY 660
DB 601 VTLANKFSSCVPRKINVSPTTYRLKDCRGVFPFRBRRELPRPFSEIIGICFLDAY 660
OY 661 QQGTNSKPCFOKKEVDGNANFLGRASGID 690
DB 661 QQGTNSKPCFOKKEVDGNANFLGRASGID 690

```

```

DB 661 QQGTNSKPCFOKKEVDGNANFLGRASGID 690
XX
XX RESULT 4
XX ABG61878
XX ID ABG61878 standard; Protein; 717 AA.
XX
XX ABG61878;
XX
XX 15-AUG-2002 (first entry)
XX
XX Prostate cancer-associated protein #79.
XX
XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
XX Mammalia.
XX
XX WO200230268-A2.
XX
XX 18-APR-2002.
XX
XX 12-OCT-2001; 2001WO-US32045.
XX
XX 13-OCT-2000; 2000US-0687576.
XX
XX 08-DEC-2000; 2000US-0733288.
XX
XX 08-DEC-2000; 2000US-0733742.
XX
XX 24-JAN-2001; 2001US-263957P.
XX
XX 16-MAR-2001; 2001US-276791P.
XX
XX 16-MAR-2001; 2001US-276888P.
XX
XX 06-APR-2001; 2001US-281922P.
XX
XX 24-APR-2001; 2001US-286214P.
XX
XX 30-APR-2001; 2001US-0847046.
XX
XX 04-MAY-2001; 2001US-288589P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Gish KC, Mack DR, Wilson KE, Afar D, Hwezel P;
XX
XX WPI; 2002-471335/50.
XX
XX N-PSDB; ABR92193.
XX
XX
XX Detecting a prostate cancer-associated transcript in a cell in a
XX patient, useful for diagnosing prostate cancer (PC) or screening
XX modulators of PC, by determining if prostate cancer-associated genes
XX are expressed in a prostate tissue
XX
XX
XX Claim 27; Page 363; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate
XX cancer-associated transcript in a cell from a patient. The method
XX comprises contacting a biological sample from the patient with
XX prostate cancer-associated polynucleotides (designated PC genes) that
XX selectively hybridise to a sequence that is at least 80% identical
XX to them. The prostate cancer-associated polynucleotide sequences
XX are differentially expressed in prostate tumour tissue or in
XX prostate cancer and are derived from the tissues of various
XX organisms such as humans or other mammals (e.g. mice, sheep and dogs).
XX The methods of the invention are useful for diagnosing and treating
XX prostate cancer in mammals. The prostate cancer-associated genes are
XX useful for diagnosing or treating prostate cancer, as well as for
XX identifying modulators of prostate cancer or agents that inhibit
XX prostate cancer. The nucleic acid sequences are particularly useful
XX in gene therapy, as a vaccine or in antisense applications.
XX ABG61800-ABG61944 represent prostate cancer-associated proteins.
XX
XX Sequence 717 AA:
SQ
Query Match 88.3%; Score 3171.5; DB 23; Length 717;
Best Local Similarity 90.7%; Pred. No. 11e-294;
Matches 625; Conservative 10; Mismatches 37; Indels 17; Gaps 4;
OY 1 MECTRLDKLKITGECFSLAPGVNPNSSSEAGSSSSCKATVPICODIPKNIQESLP 60
DB 1 MECTRLDKLKITGECFSLAPGVNPNSSSEAGSSSSCKATVPICODIPKNIQESLP 60

```

```

DB      1 MECTKTKDKITGCEPFLIAPEQVNESSEAAAGSESCATVPICQDIPKNIQESLP 60
QY      61 ORKTSRSRYVLTHTLAESICKLIPPEPERLVALQRTLAHKIKESKSLEREDPEKFTIAE 120
DB      61 ORKTSRSRYVLTHTLAESICKLIPPEPERLVALQRTLAHKIKESKSLEREDPEKFTIAE 120
QY      121 QAVAGVPEVETIKESIGEVEFKICEYEDENILGVGGTILKDFINSEFTLLKQSSHCQE-- 178
DB      121 QAVAGVPEVETIKESIGEVEFKICEYEDENILGVGGTILKDFINSEFTLLKQSSHCQE-- 178
DB      121 QAVAGVPEVETIKESIGEVEFKICEYEDENILGVGGTILKDFINSEFTLLKQSSHCQE-- 178
QY      179 -----AGKRGRLDASIIICLDKEDDLHYHYFFPKRTTSLILPGIILKAAAHVYXHEVEV 233
DB      172 QRLPSRKRGQLDASIIICLDKEDDLHYHYFFPKRTTSLILPGIILKAAAHVYXHEVEV 231
QY      234 SLMPPCFHNDCSEFVNOPIYLLSVHAKSKTSPSPKPOSSIVIPISLCKTTPPFHMD 293
DB      232 SLMPPCFHNDCSEFVNOPIYLLSVHAKSKTSPSPKPOSSIVIPISLCKTTPPFHMD 291
QY      294 KDMTIIQFNGIRRLNRRDPQCKPNEFEYFETLPRKINQTFESGIMTMLMNOFVAVRW 353
DB      292 KDMTIIQFNGIRRLNRRDPQCKPNEFEYFETLPRKINQTFESGIMTMLMNOFVAVRW 350
QY      354 DNSVKSRSRVMDLKGOMIYIVESSALIFGSPCVDRLEDPFTGGLYSLDIPHNALRDVY 413
DB      351 DNSVKSRSRVMDLKGOMIYIVESSALIFGSPCVDRLEDPFTGGLYSLDIPHNALRDVY 410
QY      414 LIGEQARADGLKRRIGLKATLEQAHQALEEKKRTVDLCSIFCEVAQOLMOGOVY 473
DB      411 LIGEQARADGLKRRIGLKATLEQAHQALEEKKRTVDLCSIFCEVAQOLMOGOVY 470
QY      474 AKKFSVNTMLFSDIVGFTALCSQSPLOYITMLNLYTRFDQCGELDYKVTETIGDAYC 533
DB      471 AKKFSVNTMLFSDIVGFTALCSQSPLOYITMLNLYTRFDQCGELDYKVTETIGDAYC 530
QY      534 VAGGLKESDTHAVOIALMAKMELEDVMSPHGPIKMRIGLHSGSVAGVYKMPR 593
DB      531 VAGGLKESDTHAVOIALMAKMELEDVMSPHGPIKMRIGLHSGSVAGVYKMPR 590
QY      594 YCLFGNNVTLANKEFSCSVPRKINVSPTTYRLKDCPGFVTPRSREELPNNPSEIPIGI 653
DB      591 YCLFGNNVTLANKEFSCSVPRKINVSPTTYRLKDCPGFVTPRSREELPNNPSEIPIGI 650
QY      654 CHEFLAYQOGTNSKPCFOKKVDGNAVF 682
DB      651 CHEFLAYQOGTNSKPCFOKKVDGNAVF 679

```

RESULT 5  
AB58042  
ID .AB58042 standard; Protein; 676 AA.

```

XX      AC      AB58042;
XX      DT      26-MAR-2002 (first entry)
XX      DE      Drosophila melanogaster polypeptide SEQ ID NO 918.
XX      KM      Drosophila; developmental biology; cell signalling; insecticide;
XX      KM      pharmaceutical.
XX      OS      Drosophila melanogaster.
XX      PN      WO200171042-A2.
XX      PD      27-SEP-2001.
XX      PF      23-MAR-2001; 2001MO-US09231.
XX      PR      23-MAR-2000; 2000US-191637P.
XX      PR      11-JUL-2000; 2000US-0614150.
XX      PA      (PEKE ) PE CORP NY.

```

```

PI      Venter JC, Adams M, Li PWD, Myers EW;
XX      WPI; 2001-656860/75.
DR      N-PSDB; AB102145.
XX      PT      New isolated nucleic acid detection reagent for detecting 1000 or more
PT      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
PS      Disclosure; SEQ ID NO 918: 21pp + Sequence Listing; English.
XX      CC      The invention relates to an isolated nucleic acid detection reagent
CC      capable of detecting 1000 or more genes from Drosophila. The invention is
CC      useful in developmental biology and in elucidating cell signalling and
CC      cell-cell interactions in higher eukaryotes for the development of
CC      insecticides, therapeutics and pharmaceutical drugs. The invention
CC      discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
CC      sequences (AB101840-AB16175) and the encoded proteins
CC      (AB57737-AB572072).
CC      The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences.
XX      SQ      Sequence 676 AA:

```

Query Match 27.6%; Score 991.5; DB 22; Length 676;  
Best Local Similarity 37.8%; Pred. No. 1,3e-85;  
Matches 247; Conservative 117; Mismatches 209; Indels 81; Gaps 20;

```

QY      58 SLPORTSRRYVLTHTLAESICKLIPPEPERLVALQRTLAH-----HKIK---ESR 106
DB      27 ALDEDEISDALTHTLHMAIQLTLTAPSNEDLTAVTSLVAKYRQWPNHKLKLPDQF 86
QY      107 KSLERDEPEKFTIAE-----QAVAGVPEVETIKESIGEVEFKICEYEDENILGVGGTIL 159
DB      87 KSCANTDYLDADIELEKDKEDASASELV-----LAGEELITCCGIIIRAFRCIGTDL 141
QY      160 KDEFLNSES---TLKQSSHCQEAQKRGRLDASIIICLDKEDDLHYHYFFPKRTTSLILP 216
DB      142 QEFLGLSDGYDVIAK-----IQEED---VYDTPVCAQGESE-----LFTSERVIMALL 189
QY      217 GITKAAAHVYXHEVEVSLMPPCFHNDCSEFVNOPIYLLSVH-----MKSTKP-SUSPSK 270
DB      190 GSKALTRMLKYDVNIKIEP--VEGDARRY---RYFLSVLYKNSOTMLGRPTYSKTI 244
QY      271 PQ-----SSIVIPISLFCCTPEPHFMDPKDMIILOFGNGIRRLNRRDPQCKPNE 320
DB      245 PETVQNSNSMASDLONNSSSFCMKPFMHETINDEQLELVOLGKFSKLY-----KPYM 297
QY      321 EE-----YFETLTPK-INQTFSGIMTMLMNOFVAVRRMDSYKKSRYMDLKGOMIY 372
DB      298 ADFGCGATYFDFKRRPRGUTMKFRDIYRTYTFEFLIGLNNPPGAVDPALGELKGMVH 357
QY      373 IVSSALIFGSCPCVDRLEDPFTGRLYSLDIPHNALRDVYLGEOARAADGLKRRIGKL 432
DB      358 CPESNLSLTFGSPFLGLDGLTNGELFISDIPHDATREVIIVGEOARAADGLRRRMDKI 417
QY      433 KATLEQAHQALEEKKRTVDLCSIFPCVDAQOLMOGOVYQAKKSVNTMLFSDIVGETA 492
DB      418 KNSIEENSAVYTERKKNVSLHLIPPAETAEKMLGSSSIDATYTPVITLFDSDIVGETS 477
QY      493 ICSQSPLOYITMLNLYTRFDQCGELDYKVTETIGDAYCVAGGLKESDTHAVOIALM 552
DB      478 ICSRAPFVIVISMLEGKDFDECDPFDYKYVETIGDAYCVAGGLHRAIYDAHKYAWM 537
QY      553 ALKAMELSDVMSPHGPIKMRIGLHSGSVFAGVYKMPRYCLFGNNVTLANKEFSCSV 612
DB      538 ALKMDIACSKHITHDEQIKMRIGLHSGVYAGVYKMPRYCLFGHSTVIANKEFSGSE 597
QY      613 PRKINVSPTTYRLKDCPGFVF--TPRSREELPNNPSEIPIGI--ICHPIDAVO 661
DB      598 ALKINVSPTTKDMITKHEGFELQRPDSFLLKRPNN--PGGTETCYFLESFR 649

```

RESULT 6  
AAVS1608  
ID AAVS1608 standard; Protein: 619 AA.  
XX  
AC AAVS1608;  
XX  
DT 26-MAR-2000 (first entry)  
XX  
DE Human soluble guanylylcyclase beta1 protein.  
XX  
KW Human: guanylylcyclase alpha1; hscGalpha1; soluble;  
KW guanylylcyclase beta1; antiarteriosclerotic; vasotrophic; hypotensive;  
KW gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;  
KW peripheral arterial occlusive disease; arterial hypertension.  
XX  
OS Homo sapiens.  
XX  
PN DEL9837015-A1.  
XX  
PD 24-FEB-2000.  
XX  
PE 14-AUG-1998; 98DE-1037015.  
XX  
PR 14-AUG-1998; 98DE-1037015.  
XX  
PA (VASO-) VASOPHARM BIOTECH GMBH & CO KG.  
XX  
PI Schmidt H, Zabel U, Poller W;  
XX  
DR WPI: 2000-184044/17.  
DR N-PSDB: AA288939.  
XX  
PT New human soluble guanylate cyclase alpha1/beta1 and the nucleic acid  
PT encoding the subunits, useful for producing diagnostic antibodies, and  
PT for somatic gene therapy of arteriosclerosis -  
XX  
PS Example 1; Page 16-17; 44pp; German.  
XX  
CC This invention describes novel purified human soluble guanylate cyclase  
CC alpha1/beta1 (hscGalpha1/beta1). The products of the invention have  
CC antiarteriosclerotic, vasotropic and hypotensive activity. Nucleic acid  
CC sequences encoding the alpha1 and/or beta1 subunit are useful for somatic  
CC gene therapy of arteriosclerosis and restenosis, ischemia (infarct),  
CC peripheral arterial occlusive disease and arterial hypertension.  
CC Antibodies to hscGalpha1/beta1 can be used for diagnosis of aberrant  
CC hscGalpha1/beta1 expression in human tissues. This sequence represents  
CC the human soluble guanylylcyclase beta1 subunit described in the method  
CC of the invention.  
XX  
SO Sequence 619 AA;  
Query Match 22.0%; Score 789.5; DB 21; Length 619;  
Best Local Similarity 35.1%; Pred. NO. 2.7e-66;  
Matches 207; Conservative 96; Mismatches 219; Indels 67; Gaps 15;  
OY 101 KIKSRKLEKEDDEKTAQVAAVAVY-----EYIKESLGEVEFKICYED-ENILG 153  
DB 30 QLDEGQFLVRIIYDDSKTYDLVAASRVLMNMGEL-QMGKFFVFCOESGYDTILR 88  
OY 154 VVGGLKLFNLSFTLKQSSHCQAGKRGLEDAISILCKE-DDELFHYVFFPKRTTS 212  
DB 89 VLGSNVRFEQLNDAL-----HDILATVYPCMRAPSEFCTAENKGLILYISREGIQ 143  
OY 213 LILPGIIKAAHVLVEFEVSLM---PCFHNDCSEFVNPYLLYSVHMKSPKPS--- 265  
DB 144 DIVGIITKVAQOIHGFTIDMKVIOQRNECDHTQ-----FLIEEKSEDEFFY 192  
OY 266 -----LSPSKPOSSLVITSLFCFTFPHFMEDKMTLLQFGNGIRLMNRDPQGRNF 320  
DB 193 EDLDRFEENGTOESNISPYT-FCKAFPHIIFDRDLVVTQGNMAYRVLPQLQ-PGNCSL 250  
OY 321 EEFYEILLPKINQFFSGIMTLMNQFVVRKMDNSVAK-----SSRVMDLKQM 370

DB 251 LSVFSLVRPHDISFHGILSHINIVFVLRSEKGLDVEKLECEDELGTETISCLRLKGM 310  
OY 371 IYIVSSAIFLFGSPCVRLDEFTGRGLYLDIPHNALRDVYLIGQBARADGKKRIG 430  
DB 311 IYLPBADSILFCSFVSANLDDLRRGILYLDIPHDTRDLVLDGEOFFREYRILTOLE 370  
OY 431 KIKATLEQAHOALEEKKRTVDLCSIFPCVAAQOLMOGVQAKRSNVTMLFSDIYGF 490  
DB 371 ILTDRQLQTLALDEDEKKTDTLLYSVLPSPVANELRRKRVRPAKRDONTVILFSGIYGF 430  
OY 491 TALGSGS-----PLQVITMNAALTRFPQCGEID---VYVETIGDAYCVAGLHESD 543  
DB 431 NAFPSKHAASGGEAMKIVLNDLTYRFDITLDSRKNPVTYVETVGGKMTVSGILPEPCI 490  
OY 544 THAVOIALMKMELSDVMSPPGEPIMRIGLHSGSVFAGVGVKMPRYCLPGNNVTL 603  
DB 491 HHASICHALDMMELIAGV-QVDSVQITIGTIGTGVYVGVIGQRRPRICLGRNVNL 549  
OY 604 ANKFESCVPRKINVSPTTYRLKDCPGVETPPRSREELPPNPFSEIIPG 652  
DB 550 TSTRFTTGKKGKINVSERYRCL-----MSPENSDPQFHLHRG 588  
RESULT 7  
ABB58149  
ID ABB58149 standard; Protein: 787 AA.  
XX  
AC ABB58149;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 1239.  
XX  
KW Drosophila: developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PERE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI: 2001-656860/75.  
DR N-PSDB: ABL02252.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure: SEQ ID NO 1239; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WPI  
CC at ftp.wipo.int/pub/published\_pct\_sequences.



QY 638 SREELP-----PMPSEIPGICHELDAVOCGINS 666  
DB 555 RENNRLPDLQTLAQLVHPLSPS-----HLLPRDLQGGCS 627

RESULT 9  
ABB59684  
ID ABB59684 standard; Protein; 1076 AA.

XX ABB59684;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster polypeptide SEQ ID NO 5844.  
XX  
XX Drosophila: developmental biology; cell signalling; insecticide;  
XX  
XX pharmaceutical.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
XX  
XX 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI, 2001-656860/75.  
XX  
XX N-PSDB; ABL03787.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX  
XX interactions -  
XX  
XX Disclosure; SEQ ID NO 5844; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
CC  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC  
CC useful in developmental biology and in elucidating cell signalling and  
CC  
CC cell-cell interactions in higher eukaryotes for the development of  
CC  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC  
CC sequences (ABB57737-ABB72072).  
CC  
CC (ABB57737-ABB72072).  
CC  
CC The sequence data for this patent did not form part of the printed  
CC  
CC specification, but was obtained in electronic format directly from WIPO  
CC  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
XX  
XX Sequence 1076 AA;  
XX

Query Match 12.9%; Score 462; DB 22; Length 1076;  
Best Local Similarity 43.6%; Pred. No. 1.7e-34;  
Matches 109; Conservative 38; Mismatches 91; Indels 12; Gaps 5;

QY 404 PIHNAIRD---VLLIGEARAQDGLKRRL---GRLKATLEQAHLSEKKKTVDLCS 456  
DB 785 PTFSTIRSNIRITKMGPCENLMDLLNMEQYANLSELVKEKTKQLSEKQRTLELLXQ 844

QY 457 IFPCVAAQOLMOGOVQAARFNSVTMLFSDIVGFTALCSQCSPLQVITMLNALYTRFDQO 516  
DB 845 ILPRVAAQOLMAGDLVEDEEFFSVITYFSDIVGFTELCARSSPMOVNFMNDLYSTFDRI 904

QY 517 CGELDYVYVETIGAYCYAAGL-HKESPTHAVQIAMLKMKELSDVMSRGEF---IK 572  
DB 905 IGFYDYVYVETIGAYCYAAGLPEPBGDKHAREIALMALDLIRAVSSFNLR-KKEYKIQ 963

QY 573 MRIGHSQSVFAGVGVAMPKRYCLGNNVTLANKFESCSVPKRVNPTTYTRLLKDCGFE 632

DB 964 IRIGHSSQSVAGVGVAKMPHICLFGDTVNTASRMESTGQDEKIHVSATFALLDKFGTF 1023  
QY 633 VFTPRSREL 642  
DB 1024 QMEGRGDVEL 1033

RESULT 10  
AAB83940  
ID AAB83940 standard; Protein; 1115 AA.

XX AAB83940;  
XX  
XX 23-JUL-2001 (first entry)  
XX  
XX Amino acid sequence of a guanylyl cyclase-2 chemoreceptor (HG-gcy-2).  
XX  
XX Soybean cyst nematode; guanylyl cyclase-1 chemoreceptor; HG-gcy-1;  
XX  
XX parasitic nematode; chemotaxis; HG-gcy-2.  
XX  
XX Heterodera glycines.  
XX  
XX WO200134791-A2.  
XX  
XX 17-MAY-2001.  
XX  
XX 02-NOV-2000; 2000WO-US30295.  
XX  
XX 05-NOV-1999; 99US-0435376.  
XX  
XX (UNNC-) UNIV NORTH CAROLINA STATE.  
XX  
XX Davis EL, Yan Y;  
XX  
XX WPI, 2001-335924/35.  
XX  
XX N-PSDB; AAF89878.  
XX  
XX Novel isolated polynucleotide encoding nematode guanylyl cyclase  
XX  
XX chemoreceptor, useful for controlling plant feeding nematodes -  
XX  
XX Example 2; Page 55-59; 69pp; English.

CC The present sequence represents soybean cyst nematode guanylyl cyclase-2  
CC  
CC chemoreceptor, designated HG-gcy-2. HG-gcy-1 polynucleotides and  
CC  
CC polypeptides are useful as active agents for controlling plant  
CC  
CC feeding nematodes. They can be used to screen for compounds that have  
CC  
CC the ability to disrupt parasitic nematode feeding or chemotaxis.  
CC  
XX  
XX Sequence 1115 AA;  
XX

Query Match 12.8%; Score 461; DB 22; Length 1115;  
Best Local Similarity 40.6%; Pred. No. 2.2e-34;  
Matches 114; Conservative 38; Mismatches 101; Indels 28; Gaps 6;

QY 376 SSAITFLGSPCVD-----RLDEFTGRGLYLSIDPIHNAIRDVYLIGEQAADGLKRR 428  
DB 812 SPALHLHLTKKDWSDSPAPRPKMETVT-----ALLQSNKTRSTNLMDHYFNM 858

QY 429 L-----GRLKATLEQAHLSEKKKTVDLCSIFPCVAAQOLMOGOVQAARFNSVTMLF 484  
DB 859 LEVYAGSLSEEEVEERTKELVEKKRTDILYRMLPKQYADKTKLQGSVEPETFDCYVVF 918

QY 485 SDIVGFTALCSQCSPLQVITMLNALYTRFDQOCEGLDYVYVETIGDAY-CVAGGLHKESD 543  
DB 919 SDVVSFTTIAKSCSPQVYNLNLNLYTLDSIIAEFDYKYETIGDYLGVSPHRRGH 978

QY 544 THAVQIAMLKMKELSDVMSRPH--GEPIMRIGHSQSVFAGVGVAKMPKRYCLFGNNV 601  
DB 979 EHAQHIAKMSLAFKMNIGSFTPIHLPIERLRLRIGIHGSGTVAGVGVSMKRYCLFGPTI 1038

QY 602 TLANKFESCSVPKRVNPTTYTRLLKDCGFEVFTPRSREL 642



CC 114,952). The protein (or variants) can be used in treatment of  
 CC natriuretic peptide disorders, and also to isolate peptides using  
 CC affinity chromatography. Antibodies with affinity for NPRB can  
 CC also be prep'd. An analogue, NPRBKC, comprising only AAs 1-433 of  
 CC the mature protein, is also claimed, as well as variant described  
 CC in AAR10867.

XX Sequence 1047 AA:

Query Match 12.8%; Score 460; DB 12; Length 1047;  
 Best Local Similarity 41.9%; Pred. No. 2.5e-34;  
 Matches 113; Conservative 39; Mismatches 76; Indels 42; Gaps 6;

QY 411 DVTLIGEDARAD-----GLKRLGK-----LKAYLEQA 439  
 DB 760 ELVILMERCMADPAERPDGQIKGIRFRNKGSTILDNLLRMEQYANLLEKVEER 819  
 QY 440 HQALEEKKKTVNDLCSIFPCVNAQOLMOGVNAKKSNTMTLFSIDVGFALCSQSCP 499  
 DB 820 TQATLEERKKAEDALYQILPHSVAEQLRGETVQAEADSVITFSIDVGFALSAESTP 879  
 QY 500 LQVITMLNALTFRDQCGELDVYKVEITGDAYCVAGGL-HRESPTHAVOJALMAKME 558  
 DB 880 MQVATLNDLDTCTCEADAIIDNFDYKVEITGDAYVWSGLPGRNGORNAPEIARMAIAL- 938  
 QY 559 LSDEYMS-----PHEPIKRIKIGLSGSVFAGVGVMPRYCLFGNNVTIANFECSV 612  
 DB 939 --DAVSSFRIRHRPH-DQLRLRIGVHTGPVCAGVGLKMPRYCLFGDVTNASHMESNGQ 995  
 QY 613 PKINVSPTTYRLKDKCGFVFTPRSREEL 642  
 DB 996 ALKIHVSSTTKDALDELGCFOLELRGDYEM 1025

# RESULT 13

AAR10867  
 ID AAR10867 standard; Protein; 1047 AA.

XX AAR10867;  
 XX  
 AC AAR10867;  
 XX  
 DT 09-APR-1991 (first entry)  
 XX  
 DE NPRB(Pro655, Glu656, Leu663, Phe664, Ala682).  
 XX  
 KW NPRB; ANP; BNP; kidney failure; heart failure; protein kinase;  
 KW hyperaldosteronism; glaucoma; guanyl cyclase.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key  
 FT Peptide 1..22 Location/Qualifiers  
 FT Protein /label- signal sequence 12  
 FT Domain /label- mature NPRB 23..455  
 FT /label- extracellular domain /note- "binds natriuretic peptides A, B and C" 456..456  
 FT Domain /label- transmembrane domain 479..1047  
 FT /label- cytoplasmic domain /note- "GC and protein kinase activity" 24..26  
 FT Modified-site /label- N-glycos- site 35..37  
 FT Modified-site /label- N-glycos- site 161..163  
 FT Modified-site /label- N-glycos- site 195..197  
 FT Modified-site /label- N-glycos- site 244..246  
 FT Modified-site /label- N-glycos- site 277..279

FT /label- N-glycos- site  
 FT Modified-site 349..351  
 FT /label- N-glycos- site 600..602  
 FT Modified-site /label- N-glycos- site

XX WO9100292-A.

XX 10-JAN-1991.

XX 22-JUN-1990; 90WO-0503586.

XX 23-JUN-1989; 89US-0370673.

XX (GETH ) GENENTECH INC.

XX Chang M, Goeddel D, Lowe D;

XX WPI; 1991-036711/05.

PT Natriuretic protein receptor B - for diagnosis and treatment of  
 PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.

PS Claim 27; Disclosure: 49pp; English.

XX The sequence is that of a variant of natriuretic peptide receptor B  
 CC (NPRB).

CC Editor's note: the AAs replaced in the sequence do not correspond  
 CC to those of NPRB given in the specification (see p.27, 1.23-27;  
 CC figs. 1 and 2), but instead to AAs in the sequence provided for  
 CC NPRB (fig. 2). The origin of the AAs to be substituted is not

CC clear; they may be taken from the sequence of rat NPRB (p.8, 1.18;  
 CC p.9, 1.33) or rat NPRB (p.27, 1.20). The substitution indicated on

CC p.9, 1.33, i.e. "Phe644" is thought to be an error and should read  
 CC "Phe664".

CC The mutant was produced by modifying the DNA encoding the  
 CC native sequence. The protein has guanyl cyclase (GC) activity and

CC protein kinase activity. The DNA can be inserted into expression  
 CC vectors for the prodn. of the NPRB analogue. The protein can be

CC used in treatment of natriuretic peptide disorders, and also to  
 CC isolate peptides using affinity chromatography. Antibodies with

CC affinity for NPRB can also be prep'd.

CC See also AAR10399.

XX Sequence 1047 AA:

Query Match 12.8%; Score 460; DB 12; Length 1047;  
 Best Local Similarity 41.9%; Pred. No. 2.5e-34;  
 Matches 113; Conservative 39; Mismatches 76; Indels 42; Gaps 6;

QY 411 DVTLIGEDARAD-----GLKRLGK-----LKAYLEQA 439  
 DB 760 ELVILMERCMADPAERPDGQIKGIRFRNKGSTILDNLLRMEQYANLLEKVEER 819  
 QY 440 HQALEEKKKTVNDLCSIFPCVNAQOLMOGVNAKKSNTMTLFSIDVGFALCSQSCP 499  
 DB 820 TQATLEERKKAEDALYQILPHSVAEQLRGETVQAEADSVITFSIDVGFALSAESTP 879  
 QY 500 LQVITMLNALTFRDQCGELDVYKVEITGDAYCVAGGL-HRESPTHAVOJALMAKME 558  
 DB 880 MQVATLNDLDTCTCEADAIIDNFDYKVEITGDAYVWSGLPGRNGORNAPEIARMAIAL- 938  
 QY 559 LSDEYMS-----PHEPIKRIKIGLSGSVFAGVGVMPRYCLFGNNVTIANFECSV 612  
 DB 939 --DAVSSFRIRHRPH-DQLRLRIGVHTGPVCAGVGLKMPRYCLFGDVTNASHMESNGQ 995  
 QY 613 PKINVSPTTYRLKDKCGFVFTPRSREEL 642  
 DB 996 ALKIHVSSTTKDALDELGCFOLELRGDYEM 1025

# RESULT 14

AAR38862





Wed Jul 2 12:21:11 2003

us-09-762-767a-2.rag

Page 12

Db 994 LFQDTVNTASRMSNGEALKIHLSSSETKAVLEEEFGGFELELRGDVEM 1040

Search completed: June 27, 2003, 13:03:00  
Job time : 39.3713 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:03:07 ; Search time 19.5034 Seconds

(without alignments)  
3401.080 Million cell updates/sec

Title: US-09-762-767a-2

Perfect score: 3593  
Sequence: 1 MFCRLKDKITGECPSFL.....QKQVEDGNANFLKASGID 690

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3332	90.0	690	1 OYRTAI	guanylate cyclase
2	3171.5	88.3	717	2 S23098	guanylate cyclase
3	3107.5	86.5	691	1 OYBO77	guanylate cyclase
4	1635	45.5	732	2 S18325	guanylate cyclase
5	976.5	27.2	683	2 JH0810	guanylate cyclase
6	795	22.1	619	1 OYRTB1	guanylate cyclase
7	793.5	22.1	619	1 OYBO70	guanylate cyclase
8	789.5	22.0	619	1 S23097	guanylate cyclase
9	774.5	21.6	682	1 OYRTB2	guanylate cyclase
10	749.5	20.9	751	2 T24458	hypothetical prote
11	587	16.3	685	2 T19968	hypothetical prote
12	582.5	16.2	686	2 T23721	hypothetical prote
13	581	16.2	1099	2 T16822	hypothetical prote
14	551.5	15.3	699	2 T18984	hypothetical prote
15	493.5	13.7	583	2 F88642	protein C54E4.3 [1
16	463.5	12.9	1057	2 T15720	hypothetical prote
17	460	12.8	1047	1 OYHUBR	natriuretic peptid
18	460	12.8	1047	1 OYRTBR	atrial natriuretic
19	460	12.8	1047	1 I45882	atrial natriuretic
20	457.5	12.7	1057	1 OYRTTR	guanylyl cyclase A
21	457.5	12.7	1057	2 I55319	natriuretic peptid
22	456.5	12.7	1061	1 OYHUBR	natriuretic peptid
23	454.5	12.6	1057	2 I57963	guanylate cyclase
24	452	12.6	1100	2 T42260	atrial natriuretic
25	449.5	12.5	1057	1 OYMSAR	guanylate cyclase
26	448	12.5	1108	2 A55915	guanylate cyclase
27	442	12.3	1108	2 JCS585	guanylate cyclase
28	440	12.2	1103	2 JCS581	guanylate cyclase
29	438	12.2	1108	2 B55915	guanylate cyclase

30	438	12.2	1110	1 S55279	guanylate cyclase
31	436	12.1	632	2 T31666	natriuretic peptid
32	435	12.1	1102	2 JH0717	guanylate cyclase
33	434	12.1	632	2 T31667	guanylate cyclase
34	430	12.0	1110	2 I59370	guanylate cyclase
35	428.5	11.9	1050	2 S45636	natriuretic peptid
36	428.5	11.9	1130	2 A89130	guanylate cyclase
37	428	11.9	1068	2 T42382	protein F52E1.4 [1
38	427	11.9	1125	1 OYURCP	guanylate cyclase
39	423.5	11.8	1525	1 A56699	speract receptor p
40	420.5	11.7	1005	2 S33525	guanylate cyclase
41	418	11.6	1012	2 T24384	hypothetical prote
42	418	11.6	1170	2 T27564	hypothetical prote
43	414	11.5	1144	2 T21223	hypothetical prote
44	412.5	11.5	1122	2 T28130	hypothetical prote
45	411.5	11.5	540	2 T34187	hypothetical prote

## ALIGNMENTS

## RESULT 1

OYRTAI

guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 chain - rat

N:Alternate names: guanylate cyclase, soluble, 77k chain

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Mar-1993 #sequence\_rev10 31-Mar-1993 #text\_change 18-Jun-1999

C:Accession: A38297

R:Nakane, M.; Arai, K.; Saheti, S.; Kuno, T.; Buechler, W.; Murad, F.

J. Biol. Chem. 265, 16841-16845, 1990

A:Title: Molecular cloning and expression of cDNAs coding for soluble guanylate cyclase

A:Reference number: A38297; MUID:91009100; PMID:1698769

A:Accession: A38297

A:Molecule type: mRNA

A:Residues: 1-690 <NA>

A:Cross-references: GB:M57405; GB:M36075; NID:q204277; PIDN:AAA41206.1; PID:q204278

A:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

C:Keywords: GMP biosynthesis; heterodimer; phosphorus-oxygen lyase

F:432-660/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match	Score	DB 1;	Length	690;
Best local Similarity	89.18;	Pred. No. 4,6e-217;		
Matches 616;	Conservative 39;	Mismatches 34;	Indels 2;	Gaps 2;
OY	1	MFCRLKDKITGECPSFLAPGVNSESSEASGSCAKAVPICODIPKNIQESLP	60	
Db	1	MFCRLKDKITGECPSFLAPGVNTEIEVAGVSECOATLPICQFAE-NAEGSHP	59	
OY	61	QRTSRVYLTHTLAEISICKLIFPERLNVALQRTIAHKIKESKSLERDEFTIAE	120	
Db	60	QRTSRVYLTHTLAEISICKLIFPERLNVALQRTIAHKIKENNSSEKEDLERIAE	119	
OY	121	QAVAGVPEVYKESGEVEFKICYEDNIIIGVGGITKIDLNSSTILKSSHQENG	180	
Db	120	EALAGVPEVYKESGEVEFKICYEDNIIIGVGGITKIDLNSSTILKSSHQENG	179	
OY	181	KGRLEDAISILCKDEDDLAHYFFPKRTSILIPGIIKAAAHVYETEVEVSLMPCE	240	
Db	180	KGRLEDAISILCKDDQDLNYYFFPKRTTALLPGIIKAAARITESHVEVSLMPCE	239	
OY	241	HNDCSEFNQPYLLSVHAKSTRPSLSPEKPOSSLVPTSLFCKTFPFHMDKMTIQQ	300	
Db	240	RECEFEVQPYLLSVHAKSTRPSLSPEKPOSSLVPTSLFCKTFPFHMDRDLATIQ	299	
OY	301	PNGIRRLRNRRDFOGKPEEFETLTPKINOTPSGIMTNMQVAVVRWDNSYKKS	360	
Db	300	LNGLIRRLRNRRDFOGKPEEFETLTPKINOTPSGIMTNMQVAVVRWDNLVYKKS	359	
OY	361	SRVMDKGMIIYVESAILFLGSPVDLEDFGTGGLYSDIPINHALRDVYLIGEQR	420	
Db	360	SRVMDKGMIIYVESAILFLGSPVDLEDFGTGGLYSDIPINHALRDVYLIGEQR	419	
OY	421	ADGGLKRLGLKATLEQAHQALEEKKRTVDLLSIFPCEVAQQLMOGVVQAKFSNV	480	



```

Db      421 RQDDDLKRLGKLTLEAHQALEEKKRTVDLLCSIPSEVAARLQMGHVAQARFEN 480
      480 VTMLFSDIYGFALISQCSPLQVITMLNLYTRPDQCGELDYKVTETGDAYCVAAGH 539
      481 VTMLFSDIYGFALISQCSPLQVITMLNLYTRPDQCGELDYKVTETGDAYCVAAGH 540
      540 RESDTHAVQIALMALKMMELSDVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFEN 599
      541 RESDTHAVQIALMALKMMELSDVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFEN 600
      600 NTLANKFESCVSPKINVSPTTYRLKCKPGFVTPKREBELPMPFSEIPIGICHLDA 659
      601 NTLANKFESCVSPKINVSPTTYRLKCKPGFVTPKREBELPMPFSEIPIGICHLDA 660
      660 YOGGTNSKRFQKRDVEDGNANFLGKASGD 690
      661 YOGGTNSKRFQKRDVEDGNANFLGKASGD 691

```

## RESULT 4

```

S18325
guanylate cyclase, soluble, alpha chain - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C/Accession: S18325
R:Harteneck, C.; Wedel, B.; Koesling, D.; Malkewitz, J.; Boehme, E.; Schultz, G.
FEBS Lett. 292, 217-222, 1991
A:Title: Molecular cloning and expression of a new alpha-subunit of soluble guanylyl cyclase.
A:Reference number: S18325; MUID:92070494; PMID:1683630
A:Accession: S18325
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-732 <HAR>
A:Cross-references: EMBL:63282; NID:931670; PID:CAA44921.1; PID:931671
C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
F:473-701/Domain: guanylate cyclase catalytic domain homology <GCC>

```

## Query Match

```

Best Local Similarity 45.5%; Score 1635; DB 2; Length 732;
Matches 355; Conservative 106; Mismatches 200; Indels 58; Gaps 15;

```

```

      13 GECPSLA-----PGVPNSESSEAAAGSESCATVPCODIPKIQESLPQRKTS 65
      27 GCPPLRLCWMGSRSPG--PLEPSPAAAAA---AAPPTPAASAAAAAATAGARRVOR 81
      66 RSRVYLHTLAESICKLIPPEFERLNAVLAORTLA--KHKT---KESKSL----- 109
      82 RRRVNLDSIGESISRLTAPSPOTIOOTLKRITQYEHQYIGYRDACKNHNISNRCSTAD 141
      110 -----ERDEFKTIABQAAVAGVPEVVKESLGEVEFKICYEDENILGVGGLDPLN 164
      142 HSNKEIEDVSGILCTANILGLKFEIQRGEFFNICFHNENHVAVAGGLTDFEN 201
      165 SFTILLKQSHQGA-GKRGRLDASILCLDKEDFLAHYYPFKTTSLILPGITKAA 223
      202 GDDALLE--HRTSFGQATLESPEFLCKELPEGLMAYPHPHIVGPFAMIGMKAAG 258
      224 HVLTEVEVSLMPCFHDND--CSEFVNOP---YLLSYHMKSTKSLSPKPOSSIVIP 278
      259 KRIYLDVVEQVA---NEKICSD-VSNPGMCSCTFLIKCEKENTINKNLPQGSQYV 313
      279 TSL-----FCKTFPHFEMFDKMTLLQFNGIRLRLMNRDPOGKPNFEYFELPKIN 332
      314 ALRLISINIFCRAFPFLHMFDEPSMSVILQEGELRKQL-RCDPHKVLKFEDECPEIVSPKN 372
      333 QFESGIMTLNMQVYVRVRMRDMSVKKSRVMDLQGMITYESSAILLGSCVRLD 392
      373 AFPERVLRLSTPEVLRTPKPEASGSSENKQVMEVKQMHPESNSILFGSCVRLD 432
      393 FTGGRGLYSIDIPINHALRDVYLIGEQARAQDGLKRLGKLTLEAHQALEEKKRTVD 452
      433 LMGRLHLSIDIPINHALRDVYLIGEQARAQDGLKRLGKLTLEAHQALEEKKRTVD 492

```

```

      453 LICSIPCEVAQOLMGOVVAQKKSNTMLFSDIYGFALISQCSPLQVITMLNLYTR 512
      493 LIXSIFPSGVAQOLMGOVVAQKKSNTMLFSDIYGFALISQCSPLQVITMLNLYTR 552
      513 FDOCGELDYKVTETGDAYCVAAGHLSKESDTHAVQIALMALKMMELSDVMSPHGEPIK 572
      553 FDOCGELDYKVTETGDAYCVAAGHLSKESDTHAVQIALMALKMMELSDVMSPHGEPIK 612
      573 MRIGHSGSVFAGVGVKMPRYCLFENNTLANKFESCVSPKINVSPTTYRLKCKPGF 632
      613 MRIGHSGSVFAGVGVKMPRYCLFENNTLANKFESCVSPKINVSPTTYRLKCKPGF 672
      633 VTPKREBELPMPFSEIPIGICHLDAVYOGGTNSKRFQKRDVEDGNAN---FLGKAS 687
      673 VTPKREBELPMPFSEIPIGICHLDAVYOGGTNSKRFQKRDVEDGNAN---FLGKAS 731

```

## RESULT 5

```

JH0810
guanylate cyclase (EC 4.6.1.2) - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 26-Aug-1999
C/Accession: JH0810
R:Toshikawa, S.; Miyamoto, I.; Aruga, J.; Furutachi, T.; Okano, H.; Mikoshiba, K.
J. Neurochem. 60, 1570-1573, 1993
A:Title: Isolation of a Drosophila gene encoding a head-specific guanylyl cyclase.
A:Reference number: JH0810; MUID:93203896; PMID:8095978
A:Accession: JH0810
A:Molecule type: mRNA
A:Residues: 1-683 <YOS>
A:Cross-references: GB:S57126; NID:q298674; PID:q298675
A:Experimental source: head
C:Genetics:
A:Gene: dgc 1
A:Cross-references: FlyBase:Fgn0010227
C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C:Keywords: GMP biosynthesis; phosphorus-oxygen lyase
F:417-641/Domain: guanylate cyclase catalytic domain homology <GCC>

```

## Query Match

```

Best Local Similarity 27.2%; Score 976.5; DB 2; Length 683;
Matches 244; Conservative 116; Mismatches 211; Indels 83; Gaps 20;

```

```

      58 SLPQRTSRSRVYLHTLAESICKLIPPEFERLNAVLAORTLA-----HKIK---ESR 106
      27 ALDELDSDALFTLHLQALQLTLAPSNEDLTATSLVAKYRQWPNHKLKLDPTF 86
      107 KSLERDEFKTIAB-----QAVAGVPEVVKESLGEVEFKICYEE-DENILGVVGT 159
      87 KSCANDYDLADIOELLKMDASASILV-----LLGELINCCCGIITERAFRCIGTDL 141
      160 KDFLNSFS---TLKQSHQGAQKRGRLDASILCLDKEDFLAHYYPFKTTSLILP 216
      142 QEFVSLDGVYDLK---LQEDD---YTDGFCVAGGE---LIFSERYVIMLL 189
      217 GITKAANLYTEVEVSLMPCFHDNDCESEFVNOPYL---YSVHMKSTKSLSPKPO 272
      190 GSKALTRMLYKVDVNIKIEPV-----ERCVAVSLPILLAGGQGLPDHADSISVSTIPE 245
      273 -----SSVITFSLFCKTFPHFEMFDKMTLLQFNGIRLRLMNRDPOGKPNFE 322
      246 TVQSRNSNNAIDLQMNSSSPCKMFPWHFTINBOLLEVLQGRFSKLY-----KPYMAD 298
      323 -----YFELMPK-INQFSGIMTLNMQVYVRVRMRDMSVKKSRVMDLQGMITY 374
      299 FGCQATTYIDFKRPKGLTKMFRDVIKRTYTPFLIGLNNPAGVDFPAGIETIGQWVHCP 358
      375 ESSAILFLGSPCYRLDFTGRLYSIDIPINHALRDVYLIGEQARAQDGLKRLGKLT 434
      359 ESNLSLFLGSPCYRLDFTGRLYSIDIPINHALRDVYLIGEQARAQDGLKRLGKLT 418
      435 TLEQAHQALEEKKRTVDLLCSIPCEVAQOLMGOVVAQKKSNTMLFSDIYGFALIS 494

```

Db 419 SIEEANSVTERKKKNVSLHLHPAEIAEKLKMGSSIDAATYPPDVLTFSDIYGFISIC 478

Qy 495 SQCSPLQYITMLNLYTFEDDQCGGELDYKXVETIGDATCAVAGLHKES--DTHAVQILM 552

Db 479 SRAPFPMVISMEGLYKDFECDFDYKXVETIGDACVAGSLHRAISYAHNR--CLD 535

Qy 553 ALKMMELSDYMSPHGEPICKMRIGLHSGSVAGVGVMPRCYCLGNNVTLANKFESCSV 612

Db 536 GLKMDACSKRITHDGEQIKMRIGLHTGTVLAVGGRMPRCYCLGHSVTIANFESSE 595

Qy 613 PRKINVSSTYRNLKDCGFYF--TPRSRELLPRPFSEIIG---ICGFDAIQ 661

Db 596 ALKINVSSTYTDMLTKHGEFEELQPRDPSFLPKEFPN--PGGVTCTCLFSFR 647

## RESULT 6

guanylate cyclase (EC 4.6.1.2), soluble, beta-1 chain - rat  
N:Alternate names: guanylate cyclase, soluble, 70K chain  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1993 #sequence-revision 31-Mar-1993 #text-change 18-Jun-1999  
C:Accession: A31871  
R:Nakane, M.; Saheki, S.; Kuno, T.; Ishii, K.; Murad, F.  
Biochem. Biophys. Res. Commun. 157, 1139-1147, 1988.  
A:Title: Molecular cloning of a cDNA coding for 70 kilodalton subunit of soluble guanylate  
A:Reference number: A31871; MUID:89087429; PMID:2905128  
A:Accession: A31871  
A:Molecule type: mRNA  
A:Residues: 1-619 <NNA>  
A:Cross-references: GH:M2562; NID:g204273; PIDN:AAA41204.1; PID:g204274  
C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology  
C:Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase  
F:373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match	22.18;	Score 795;	DB 1;	Length 619;

Best Local Similarity 34.3%; Pred. NO. 1.9e-47;  
Matches 210; Conservative 100; Mismatches 225; Indels 78; Gaps 15

```

09      LNVALORTLAKHKINESRKSLERE-----DFEYT---IAEQAVAGVYE 13G
89
Db      5  VNAHELIVIRNYGPEWEMEDIKREALDOEGFLVRIYDHKSTYDVAASAKILNLNG 64
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

131 VIKESLGEVFKICYED-ENILGVVGCTLDFLNSEFTLLKQSSHCOEAGKRGRLDAS 189

0Y 190 IICIDKE-DDFLHYVYFFPKRTTSLILPGIITKAAAHVLYETEVEVSLM-----PPCFHNDG 244

```

0Y      245 SEPNOPIIYSVHMKSTBS-----LSPSKPOSSLVPTSLFCTFPFHMFDDKM 296
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

297 TILQFNGJIRLMNRDFOGKPNFEEYFEILTPKINOTSGINTMLNQFVYVRMDNS 356

DB 228 VVICGNAIKVLPDQ - PGCSLSVSFLVRPHIDISFHLSHINVFVLSKEGLD 286

QY 357 VKK-----SSRYMDLKQMIYIVESSAILFQSPCVURLLEDPTGRGLYLDPIH 406

Db 287 VERKECEDELTGAETISCLRLKGQMIYLPKADSIILFLCSPSYNNLDDLTRGLYLSDIPLH 346

467 WOGGVOVAAKREESNTMI ESTVJCEMA ICSOGS-----DI QUTMI NAY VMBREOOOOOCCER D- E33  
468  
469  
470

Db 407 RHKRPPVPAKRYDNTILFSGIIVGFNPFCSKTIASGEGAMKIYNLLNDLITRFIDILTDSRKN 406

Db 467 PEYKKEVETADKMTVSGSLPEPCIHARSICHLADMMELINGOV-QYDSESVQTTIICHT 525

QY 580 GSEFAGVGVCKMRGRCYCLGNNTYLNKFECSQVPRKINSPTYYALDCCGFPTRRSR 639

Db 526 GEVYTVIGVGRMRCLTGFNTVNLISRTTETGEKGINVSEYTKCL-----MSP 575

QY 640 EELPPNPFSEIPG 652

Db 576 ENSDPQFHLEHRG 588

## RESULT 7

guanylate cyclase (EC 4.6.1.2), soluble, beta-1 chain - bovine  
M.Alternate names: guanylate cyclase, soluble, 70K chain  
C.Species: Bos primigenius taurus (cattle)  
C.Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 18-Jun-1999  
C.Accession: S01653; A38759  
R.Kossling, D.; Herz, J.; Gausepohl, H.; Nitromand, F.; Hansch, K.D.; Muejsch, A.; Bo  
FEBS Lett. 239, 29-34, 1988  
A.Title: The primary structure of the 70 kDa subunit of bovine soluble guanylate cycl  
A.Reference number: S01653; MUID:89031214; PMID:2903071  
A.Accession: S01653  
A.Molecule type: mRNA  
A.Residues: 1-619 <KOE>  
A.Cross-references: EMBL:Y00770; NID:9407; PIDN:CAA68739.1; PID:9408  
A.Accession: A38759  
A.Molecule type: Protein  
A.Residues: 1-6;28-32;41-45;96-109;337-343;396-406;562-569 <KOE2>  
C.Genetics:  
A.Introns: 281/3; 392/2  
C.Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology  
C.Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase  
F:373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match	22.18;	Score 793.5;	DB 1;	Length 619;
-------------	--------	--------------	-------	-------------

Best Local Similarity 35.3%; Pred. No. 2.5e-47;  
Matches 208; Conservative 97; Mismatches 217; Indels 67; Gaps 16

QY 101 KIKESRKSLEDEFEKTIABQVAAGVPV-----EYIKESIGEEAFKICYED-ENILG 153

Qy 154 VVGGLIKDPIINSFSTLLKROSHCOEAGKRGRLEDASILCLDKR-DDFLHVVYFFPKRTTS 2122  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 80 WCGSPGATCWDAI--SSNLTWVGVGVADPMDMMAISGVTITLGGGSGS 1422

213 LILPILIKAAHVLTEVEVSLM-----PCPFHNDOSFVNPYLLYSVHKSTKPS--- 265

QY 266 -----LSPKPOSSLVIPTSLCKTFFPHFEMEDKMTILQFNGIRRLMNRDFOGKPNF 320

00 193 EDDUNE EENNOI QESK1SFI1 FCAAE FPH1LF DRDLVAI DCGM1TRVL P0D4 -PGMCSL 4500

01 321 EEFYEILTPRINQTFSGIMTLMNMGFVVRVRWMDNSYKSS-----RVMDLKGQM 3700

Db 25. LSVESLVRPHIDISPHGILSHINTEVFLRSKEGLDVERSECEDELTGTETISCLRKGQM 310

Qy 371 IYIVESSAILFLGSPCVRLLEDFTFGNGLYLSDIPHNALRDVYLIGEQARADGLKRRLG 430

Db 311 IYPEADSIIFLCSSVMNDDLTRGLYSIDIPLDHATRDVLVLGGQFREERYKLTQELE 370

Db 371 ILFDRIQLTFLRALDEDEKKKKDTLLYSVLPSVANETLRHKRPVPAKRYDNWTLIFSGITVF 430

Db 431 NAFCSKHASGEGAMKIVNLNDLYIRFDLTDSRKNPFTYIKVETVGDKYMTYSGLP EPCI 490

Db 491 HHASICHALDAMEIAGV-QVDSYQITIGITGEVYGVICQRPRLPGNTVNL 549  
 QY 604 ANKESCSVPKRNINSPPTTYRLKDCPGFVTPRSREELPPNPFSEIPG 652  
 Db 550 TSRTETGKGNINSEYTRCL-----MTP---ENSDPQFLEHNG 588

## RESULT 8

S23097

guanylate cyclase (EC 4.6.1.2), soluble, 70k chain - human

C:Species: Homo sapiens (man)

C&gt;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 20-Apr-2000

C:Accession: S23097

R:Giull, G.; Scholl, U.; Bull, F.; Guellaen, G.

FEBS Lett. 304, 83-88, 1992

A:Title: Molecular cloning of the cDNAs coding for the two subunits of soluble guanylyl

A:Reference number: S23097; MUID:92316204; PMID:1352257

A:Accession: S23097

A:Molecule type: mRNA

A:Residues: 1-619 &lt;GIV&gt;

A:Cross-references: EMBL:X66533; NID:g31685; PIDN:CAA47144.1; PID:g31686

C:Gene: GDB:GDCY1B3; GUC1B3; GC-SB3

A:Cross-references: GDB:141992; OMIM:139397

A:Map position: 4q31.3-4q33

C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

C:Keywords: heterodimer; phosphorus-oxygen lyase

F:373-607/Domain: guanylate cyclase catalytic domain homology &lt;GCC&gt;

## Query Match

Best Local Similarity 22.0%; Score 789.5; DB 1; Length 619;

Matches 207; Conservative 96; Mismatches 219; Indels 67; Gaps 15;

QY 101 KIKESRKSLEDEFEKTAEOAAGVAVP-----EVIKESIGEEVFKICYED-ENILG 153  
 Db 30 QLEDEGQPLVRIYDSTYDVAASKVLMNAGEIL-QMGKRFVFCQSGDITLR 88  
 QY 154 VVGTLKQPLNSFTLLKQSSHCQAGKRLDASIIICLDE-DDELHYVFPFKRTTS 212  
 Db 89 VLGSVAVREFLQNLDA-----HDHLATYFGKRAPSFCTDAEKGLIHYSRERGLQ 143  
 QY 213 LILPGLIAAAHVLYETFEVSLM-----PFCFHNCSSEVNPVILLYSHMSTPS--- 265  
 Db 144 DIVIGITVMOQIGTEIDMKAVIQORNECDHTO-----FLIEEKSEEDFY 192  
 QY 266 -----LSPSKPOSSLVPTSLFCKTFPFHFMEDKMTILQFNGIRLNNRDFQKPF 320  
 Db 193 EDLDREENGQESRISPYT-FCKAFPHILIEDRLVYQCGNATYRVLPQIQ-PGNCSL 250  
 QY 321 EEPFILLPKINOTFSGITMLNMQFVAVRVRRDMSVKK-----SSRYMDLKGOM 370  
 Db 231 LSVFSLVRPHIDISPHILSHINTYFVLSKSGELDVEKLECEDELGTETISCLRLKGM 310  
 QY 371 IYIVSSALFLFGSCVRLDEFTGRGLYLSIPILHNLADVLIGEOARADGLKRRIG 430  
 Db 311 IYLPADSLILFSCSVANLDDLTGRGLSLDIPLDATRDVLVLEQOREEYKTLQLE 370  
 QY 431 KLKATLEOAHOLEEKKRTVDLCSIFPCVAAQOLMOGOVQAOKKFSVNTMLFSDYGF 490  
 Db 371 ILTDLQTLRLALEDEKKKTDTLLYSVLPVSANELRHRPAPARRYDVTLLFSGYGF 430  
 QY 491 TALISQCS-----PLQVITMLNALYTRPDQCGELD---YKVEITGDAYCVAGLHKESD 543  
 Db 431 NAFCSKHAISGEGAMKIVMLNDLYTRFPLTDSRRKNPFYKKEITGDKMTVSGLEPC 490  
 QY 544 THAVQIALMALKMLLSDVMSPHGEPIKRIGLHSGSVFAGVGVKMPRYCLFGNVT 603  
 Db 491 HHASICHALDAMEIAGV-QVDSYQITIGITGEVYGVICQRPRLPGNTVNL 549  
 QY 604 ANKESCSVPKRNINSPPTTYRLKDCPGFVTPRSREELPPNPFSEIPG 652  
 Db 550 TSRTETGKGNINSEYTRCL-----MTP---ENSDPQFLEHNG 588

## RESULT 9

OYRTB2

guanylate cyclase (EC 4.6.1.2), soluble, beta-2 chain - rat

N:Alternate names: guanylate cyclase, soluble, 76k chain

C:Species: Rattus norvegicus (Norway rat)

C&gt;Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Jun-1999

C:Accession: A36228

R:Yuen, P.S.T.; Potter, L.R.; Garbers, D.L.

Biochemistry 29, 10872-10878, 1990

A:Title: A new form of guanylyl cyclase is preferentially expressed in rat kidney.

A:Reference number: A36228; MUID:91105012; PMID:1980215

A:Accession: A36228

A:Molecule type: mRNA

A:Residues: 1-682 &lt;YUE&gt;

A:Cross-references: GB:M57507; GB:J05308; NID:g204279; PIDN:AAA41207.1; PID:g204280

A:Experimental source: kidney

C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

C:Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase

F:360-584/Domain: guanylate cyclase catalytic domain homology &lt;GCC&gt;

## Query Match

Best Local Similarity 21.6%; Score 774.5; DB 1; Length 682;

Matches 190; Conservative 107; Mismatches 204; Indels 81; Gaps 10;

QY 129 VEVIKESIGEEVFKICYED-ENILGVVGTLLKDFLNSFTLLKQSSHCQAGKRLD 187  
 Db 1 MEALIKLFEYEFKCKCKSGYDRLRTLGNTLFTENLDA-----HSYLAISQENMA 55  
 QY 188 ASILCLDKDDDLHYVFPFKRTTSILPLGIIKAAHVLYETFEVSLMPPCFHNDCEE 247  
 Db 56 PSFVRVEEGADGAMLLYVSDRGHGLHYGIIIEAVALKDFPDVMSIL-----DMNEE 109  
 QY 248 V-----NOPYLTSYVHKSTKP--SLSPKPOSS----- 274  
 Db 110 VERTGKKEHVAVLVYQAKHQRIGAKASNPQSSQADQELQGLLMKRYNTLPYC 169  
 QY 275 -----LVIPISLCKTFPFHFMEDKMTILQF 302  
 Db 170 PGEKSHSTAVRASVLEFGKPLNDTFQVYPERLWVEEVEFCAPPHIYFDEALRYKQG 229  
 QY 303 NQIRRLMRNRDQGRNPEFEYFELPKINOTFSGITMLNMQFVAVRVRR-NDNSYKKS 361  
 Db 230 VNIQKVPGLIQ--FFALDEYFSIIHPQTFNISSICKFINSQFVLTKEKMPKRAKQQ 288  
 QY 362 RYMDLKGOMIYVESALFLFGSCVRLDEFTGRGLYLSIPILHNLADVLIGEOARA 421  
 Db 289 PMLKIRGOMIMHESLRCLMFCSPNVRSLQLEESKMHSLDAPHDITDILLNQRLA 348  
 QY 422 QDGLKRLGLKATLEOAHOLEEKKRTVDLCSIFPCVAAQOLMOGOVQAOKKFSVNT 481  
 Db 349 EMELSQLEKKKEELRVLSNHLAIEKKKETILLYALPVRVANOQLEKGRVAAAGEFCT 408  
 QY 482 MFSDFVGTALCSQCSPPQVITMLNATLRPDQCGELDYYKVEITGDAYCVAGLHKE 541  
 Db 409 ILFSDVYFTNTCAACEPIQIVNMLNSMSKDRILTSVDYVKVETIGAYVVGVPVP 468  
 QY 542 SDTHAVQIALMALKMLLSDVMSPHGEPIKRIGLHSGSVFAGVGVKMPRYCLFGN 600  
 Db 469 VESHQVRANFALGMRISAKVEYRNPTGPIQIRGIRHGPVLGAGVGMKMRVCLFGT 528  
 QY 601 VTLANKFESCSVPKRNINSPPTTYRLKDCPGFVTPRSREEL 642  
 Db 529 VMTASRMESHGIPSKVHLSPYLRALKN-KGFEIYVRGEIEV 569  
 RESULT 10  
 T24458  
 hypothetical protein T04D3.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C:Accession: T24458  
 R:Kershaw, J.



submitted to the EMBL Data Library, October 1996

A:Reference number: Z19893

A:Accession: T24458

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-751 <WIL>

A:Cross-references: EMBL:Z81114; PIDN:CA803288.1; GSPDB:GN00019; CESP:T04D3.4

A:Experimental source: clone T04D3

C:Genetics:

A:Gene: CESP:T04D3.4

A:Map position: 1

A:Insertions: 1/3; 26/2; 95/3; 149/1; 213/3; 247/3; 291/3; 332/3; 369/2; 418/1; 462/3; 498/

C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

Query Match

Best Local Similarity 20.9%; Score 749.5; DB 2; Length 751;

Matches 192; Conservative 125; Mismatches 239; Indels 53; Gaps 16;

```

QY 116 KTAAQVAAVAVPEVKESEGEVEFKICEED-ENILGVGGTLDKPLNSFTLKOSS 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 50 RLNSMANVIGIPLEEIMVAGFLIQFTMETGDELRAAPDLDEGLDLSLHAFID 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 175 HCEAGRGRLDASILCKEDDEFLHYVEFFPKRTSLIPGIRKAAHVLVEYEVS 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 H---VYKTKLRGPFSCDVQADCTLLHHTYKRSGLPIYKGVREYARRITDEYVMK 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 LMP-PCFHNDSEFVNOPYLLYSV---HMKSTKPSLSPKPOSLVPTSL-----F 282
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 VOERKQHLDA--FVTE-HVFEVITQIENANSSTPKSISSKADQIDLSGIYEISSDF 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 283 CKRPFFHMDKMTILIQPNGIRLRANRDFOGKPFEEFELLTKINOTSGIWTML 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 224 SLAPFHYICDPDLFVHFENFKITPPNA-MKQETVTLLELVHPEVPSSESIXYK 282
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 343 NMOFVAVRRWDSVSKS---SRVMDLKGMIYVESAILFLGSPCVDRLEDFTRGLY 399
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 NSLFEFLKGLGDIHNADEAKVLLKGMVFIIDEKYLIMQSVAVTVRELIERNH 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 400 LSDPIFNALRDVYLIGEOARADGKRLKGLKATLEQAHOLEBEKKKTVDLCSIFP 459
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 343 LSPMRIDGTRDVIYMLNOSMSQVELNRTLEETTKIKKMAOELEIKQTDLECLMP 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 460 CEVAQOLMOGOVQAKFSNVTMLFSDIVGFTALICSCSPLOYTMNALYTRDOCGE 519
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 ASVADSLRSKADAEFACTLEFDIVFTNCACTPDVATLNDLYLRDRVGL 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 520 LDVYKETTIGDAYCVAGLHRESPTHAVOIALMALKMMLSDVMSF--HGEPIKRIGL 577
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 463 HDAYKVTETIGDAYMIVGVPERCENHAEVRLNISIGLMESKLVLSPIFH-KPIKIRLGV 521
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 578 HSGSVFAGVGVAMPKRYCLFGNNVTLANKEFSCSVPRKINVSPT-TYRLKDCPGFVTP 636
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 522 HGPVYVAGVYGIKPRYCLFGDVTWANKMESNGIOCKIHVSEGLKGLKANSYFID 581
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 637 RSREE-----LPPNFPSEIRPGI-----HFLDAYOOGTNSKCFQKK----- 673
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 KGTETLNGKAMATTYFLERNDKSWELCSRPNSGEQFTDGYME-LHDQSYIEGEGQOE 640
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 674 --DVEDGNA 680
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 641 NLTVENGNS 649

```

# RESULT 11

T19968

hypothetical protein C46E1.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T19968; T23092

R:Lennard, N.

submitted to the EMBL Data Library, November 1997

A:Reference number: Z19205

A:Accession: T19968

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-685 <WIL>

A:Cross-references: EMBL:AL008867; PIDN:CA15512.1; GSPDB:GN00028; CESP:C46E1.2

A:Experimental source: clone C46E1

R:Lennard, N.

submitted to the EMBL Data Library, October 1997

A:Reference number: Z19673

A:Accession: T23092

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-685 <WIL>

A:Cross-references: EMBL:Z99942; PIDN:CA17073.1; GSPDB:GN00028; CESP:C46E1.2

A:Experimental source: clone H13N06

C:Genetics:

A:Gene: CESP:C46E1.2

A:Map position: X

A:Insertions: 11/3; 36/2; 70/1; 133/2; 179/1; 202/3; 258/2; 298/2; 301/3; 325/2; 405/2;

C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

Query Match

Best Local Similarity 16.3%; Score 587; DB 2; Length 685;

Matches 185; Conservative 114; Mismatches 254; Indels 106; Gaps 20;

```

QY 107 KSLERDEFEKTIAEQAV-----AAGVPVEIKESLGE--EVFKICEED-- 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 KVLERAFEFAG-KENIINHYSDADFEVLDAVSILKREQVMEYGCFLIQYMETG 92
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 149 -ENILGVGGTLDKPLNSFTLKOSSH--QEGKRGRLDASILCKEDDEFLHYVEFF 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 WDLINSNPLNGFLDNLDSLHFIIDHVVYKANKLR---PFRCDNDPGTITLHYTF 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 PKRTSLILPGIIRKAAHVLVEYEVS--MPFCFHNDSEFVNOPYLLYSVHM--- 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 GRGLYIVAGVIRAEAKRFKLDVSWITIGTRORSVQATGRIIE-HVIFVKTMTD 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 KSTKPSLSPKPOSLVPTSL---PCKFPFHMDKMTILQNGNIRLRANR--RD 313
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 208 QSNDEALGIVAVOHSNNYKIRLTHMDISTEPYHAMVDDCKIVQV--RELNYNHPK 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 F--QKRPFEVEFELLPKINOTFSGIMTMNQFVVRVRMD---NSVKSSRYMD- 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 265 LLSVGRP-LMRIFEVTRPQIPIDPDSICNINAVFVQVYTTTMEORANKRAAQITA 323
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 366 -----LKGMIYVESAILFLGSPCVDRLEDFTRGLYLDIP 404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 324 SENLYEDNNGALALSGOHLKLGOMMLMSGGHIMYLCSPYTSIPELLQYGLRLAMP 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 405 IHNALRQVYLIGDQARADGKRLKGLKATLEQAHOLEBEKKKTVDLCSIFPCVAG 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 384 IHDPTRLILLNQORLSDVEMNLQLEANNDELMMAKDLVEVGKTDALLREMLPSVAG 443
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 QLMQGVQAKKFSNVTMLFSDIVGFTALICSCSPLOYTMNALYTRFOOGCELDYK 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 444 QLMQGLSVKAREBEATVMTDPTFOQIPLCTPKADIVALLMELTPKPRRLIGIQAKY 503
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 525 VETIGDAYCVAGLHRESPTHAVOIALMALKMMLSDVMSF--HGEPIKRIGLHSGV 583
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 504 VETVGDYSMSVGGIPDLVDHCEVICHIALGVMEARVCDPTNPLHIRAGIHSGPV 563
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 584 AGVYGVAMPKRYCLFGNNVTLANKEFSCSVPRKINVSPTTYRLKDCPGFVTPRSRE-- 641
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 564 AGVYGVAMPKRYCLFGDVTWNSRSHSPGRHICSENAKCAESTGREFEPRGRVQIR 623
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 642 -----LPPNFPSEIRPGI-----CHFLDAYOQ-----GTNSKPC 669
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 GKEMATYFLLRKFSKSIWEIIDRRDENCNSIDGYNELREGVYDVLYANKVYQKSKTC 682

```

# RESULT 12

T23721

hypothetical protein M04G12.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans





A:Residues: 1-699 <WIL>  
 A:Cross-references: EMBL:Z77652; PIDN:CA801118.1; GSPDB:GN00023; CESP:C06B3.8  
 A:Experimental source: Clone C06B3  
 C:Genetics:  
 A:Gene: CESP:C06B3.8  
 A:Map position: 5  
 A:Insertions: 1/3; 60/1.192/3; 210/2; 270/2; 336/2; 415/2; 659/1  
 C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

Query Match 15.3%; Score 551.5; DB 2; Length 699;  
 Best Local Similarity 26.6%; Pred. No. 2.1e-30;  
 Matches 153; Conservative 112; Mismatches 240; Indels 71; Gaps 12;

```

QY 129 VEYKESLGEV--FKICEED--ENILGVGGLDPLNSFSTLLKQSSH--COEAKR 182
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 LKATKQDVMEMVGGFLLTYSMEIGMDLVSMSPNLKGLDNLSDHYFDHYVYKANLR 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 183 GRLEDAIILCLDKEDDLHYVFPFKRTSLILPGIITAAAHVLEYEVESLMP--PPC 239
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 G---PFRCEEPDGTLLAHFTGRGLHYIVGVEAKRVEDIDITLVVGRTQRS 174
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 FHNDCSEFVNO-----PYLXSVHKSTKPSLSPKPOSSLIPTSL---- 281
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 VHNNGRVEEHVYFLINKGGOEHPLI-AIVVTGSEPRRDSGSEVSLTSTNANPT 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 282 -----FCKTFPHFMDKMTILQFNGIRRLNRRDQGRPNFEYFEILTP 329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 IVDITLISLDDFSKALPYHFVIDESCKLVCGSELNHNIPNELLOPGLIRIFEINRP 293
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 KINOTFGIMTMNOFVVRV-----RWDNSVKRSSNVMO----- 365
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 QILPDEFNIGNFNAIVAVLOVKTSPKKHMDAMSQELKQEMTTLEDATNELTQGHHL 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 366 -LKGOMIYESSAILEFLGSPCYDRLEDFTGRGLYSLDIPINALRDVVLIGEQAQDG 424
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 KLKGOMMLASKRIIILYCSPPVYSINELMQYGRMLAMPDHATRLDILLNQRLDVE 413
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 425 LKRLGLKATLRBAHALBEKKTYDLICSTFPCVAAQOLMOGOVQAQKSSNVTMLF 484
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 VNLQLEANNQLEMTLELEKRTSLIKDMLPRIPAAOOLSGEHIACE-HEAVYMF 472
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 485 SDIYGFALISCSPLQVITMLNLYRFPQOCGELVYKVFETIGDAYCVAAGLHKESDT 544
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 473 CDLPAPQOALPQCSPKRIYVMLNFIKRLDRIVIRGVYKVEYSDBTMAVSGIPDTPE 532
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 545 HAVOIALMALKMKELSDVNSPHGE-PIKMRIGLSGSVFAVGVYKMPRYCLFGNNVTL 603
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 533 HAENMCHVALGMMMEARSVIDPVSKTFLRLRIGHSCTITAGVGVTHPRYCLFGEVTL 592
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 604 ANKFESGVSVPKRTINVSPTTYRLKDCPGVFYTPRSR 639
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 593 ASQMESLGMAGKIQCSKMAVQAKMETGRFEFSPRGR 628
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 15

F88642  
 Protein C54E4.3 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: F88642  
 R:anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: F88642  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-583 <STO>  
 A:Cross-references: GB:chr\_IV; PIDN:AB92031.1; PID:g2702405; GSPDB:GN00022; CESP:C54E4.  
 A:Note: Similar to guanylate cyclase  
 C:Genetics:

A:Gene: C54E4.3  
 A:Map position: 4

Query Match 13.7%; Score 493.5; DB 2; Length 583;  
 Best Local Similarity 25.7%; Pred. No. 1.8e-26;  
 Matches 139; Conservative 129; Mismatches 218; Indels 55; Gaps 13;

```

QY 122 AVAAGVVEVYKESIGE-EYFKICEEDENILGVGGLDPLNSFSTLLKQSSHCOEAG 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 50 ATYVGLSDVDMEMVGGFLLTHACETGQMQLPCGMANNLODFDNLNSM--HYFDQIA 106
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 KGRLEDAIILCLDKEDDLHYVFPFKRTSLILPGIITAAAHVLEYEVESLMP--PPC 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 FKSEMKGPFOCEPFEGSGGLKHYFSFROGLPIYVKLAVKRTATLPEMDVQKMLERNQ 166
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 HNDCESEVNOPYLIXSVH-----MKSTK-----PSLSKPOSSLIPTSL 281
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 ERKSGMVE--HATFSEPDNNHKKGRRLPHKFRNTTETNAPSFTLS--STLVGLRD 221
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 282 FCKTFPHFMDKMTILQFNGIR--LMNRDFOGRPNFEYFEILTPK-INOTFSG 337
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 222 FKNIFPVHVCENKOMITEHIGIYLLREYGLENKTKL-----VSDLMQLVOPSDIQLYKN 277
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 338 IMYMLNQFYVRRRND--NSYKSS-----RYMDLKGMITYESSAILEFLGSPCYDR 390
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 278 VLSYLTNLTFTFOLKHSKRNEVOGSESEAPQOPLVYKGEKMPINDGNSIIFICSPIHTYV 337
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 391 EDFTRGLVYSDIPINHALRDVVLIGEQAQADQLKRLGLKATLEBAHQALEEEKKT 450
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 338 RDLINLKLITSDMPHMDATDLYMLNQRSLCQHELNKKELETKKKKKMELEEVKKSOT 397
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 451 VDLICSTFPCVAAQOLMOGOVQA-----KKSNTVTLFSDIVGFTAICSCSP 499
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 DRLTFEVPVPIAEALRAATVPALMQRISIDFSEFSQCSYIFDIPDFTISVNCSP 457
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 500 LQYTMNALYTRPDQCGELDYKVFETIGDAYCVAAGLHKESDTNAVOIALMALKMKEL 559
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 458 TEITVTVDLFRHRDRIERHKGYKVLSDVSLVGVGNANQYHCEDSLALGLLFE 517
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 560 SDEVMSPHGE-PIKMRIGLSGSVFAVGVYKMPRYCLFGNNVTLANKFESGVSVPKRTIN 618
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 518 AKQVYVPLKRSVRLRIGVAGCPVAGIVSQAPRQCVLGNVTNVTIKSHSPGKVLV 577
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 619 S 619
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 578 S 578
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: June 27, 2003, 13:05:23  
 Job time : 21.5034 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:00:58 ; Search time 34.7899 Seconds  
(without alignments)  
4086.601 Million cell updates/sec

Title: US-09-762-767a-2

Perfect score: 3593  
Sequence: 1 MFTKRLDKLTGECPSFL.....OKKVEDGNANFLGRASGID 690

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SP\_ARCHAEA:\*
- 2: SP\_BACTERIA:\*
- 3: SP\_FUNGI:\*
- 4: SP\_HUMAN:\*
- 5: SP\_INVERTEBRATE:\*
- 6: SP\_MAMMAL:\*
- 7: SP\_MNC:\*
- 8: SP\_ORGANELLE:\*
- 9: SP\_PHAGE:\*
- 10: SP\_PLANT:\*
- 11: SP\_PROTOZOA:\*
- 12: SP\_VIRUS:\*
- 13: SP\_VERTEBRATE:\*
- 14: SP\_UNCLASSIFIED:\*
- 15: SP\_VIRUS:\*
- 16: SP\_BACTERIAP:\*
- 17: SP\_ARCHAEA:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3587	99.8	690	4	08TAH3	08tah3 homo sapien
2	3266.5	90.9	691	11	09DBQ3	09dbq3 mus musculu
3	3265.5	90.9	691	11	09ERL9	09erl9 mus musculu
4	2954	82.2	564	4	09NNW8	09nnw8 homo sapien
5	2014.5	55.9	678	13	09PW12	09pw12 oryzias lat
6	2009.5	55.9	678	13	P79997	P79997 oryzias lat
7	2005	55.8	675	13	09OVV5	09ovv5 fugu rubrip
8	1097	30.5	699	5	077105	077105 manduca sex
9	991.5	27.6	676	5	024085	024085 drosophila
10	983.5	27.4	676	5	0958Q4	0958q4 drosophila
11	823	22.9	620	11	08R5L4	08r5l4 rattus norv
12	788	21.9	617	13	09OVY5	09ovy5 fugu rubrip
13	785.5	21.9	600	5	077106	077106 manduca sex
14	784.5	21.8	742	11	091XJ7	091xj7 rattus norv
15	781.5	21.8	620	11	054865	054865 mus musculu
16	781	21.7	614	13	P79998	P79998 oryzias lat

17	771.5	21.5	636	11	0920Q1	0920q1 rattus norv
18	750.5	20.9	604	5	095NK5	095nk5 hemientrot
19	749.5	20.9	751	5	002298	002298 caenorhabd1
20	743	20.7	758	5	024086	024086 drosophila
21	743	20.7	787	5	09VA09	09va09 drosophila
22	720.5	20.1	649	5	017010	017010 anopheles g
23	661	18.4	940	5	076340	076340 manduca sex
24	642.5	17.9	690	5	09VFC5	09vfc5 drosophila
25	587	16.3	685	5	09XTE0	09xte0 caenorhabd1
26	582.5	16.2	686	5	P92006	P92006 caenorhabd1
27	555	15.4	752	5	022301	022301 caenorhabd1
28	551.5	15.3	699	5	022301	022301 caenorhabd1
29	547	15.2	684	5	09B180	09b180 caenorhabd1
30	493.5	13.7	583	5	044468	044468 caenorhabd1
31	479	13.3	1047	13	09PMH0	09pmh0 xenopus lae
32	478	13.3	1276	5	08T8L6	08t8l6 caenorhabd1
33	477	13.3	1055	13	09YGW3	09ygw3 oryzias lat
34	473	13.2	1055	13	09B0U1	09b0u1 oryzias lat
35	463.5	12.9	1057	5	018331	018331 caenorhabd1
36	462	12.9	1076	5	09WZP1	09wzp1 drosophila
37	461	12.8	1112	5	09GR36	09gr36 heterodera
38	460	12.8	344	11	08R1P8	08r1p8 mus musculu
39	460	12.8	618	4	08TA93	08ta93 homo sapien
40	456.5	12.7	1172	5	09V079	09v079 drosophila
41	455.5	12.7	1056	13	09Y117	09y117 squalus aca
42	454.5	12.6	1057	11	091X04	091x04 mus musculu
43	453.5	12.6	210	13	09NH71	09nh71 epratretus
44	453	12.6	1076	5	09VP19	09vp19 drosophila
45	452	12.6	1100	11	054884	054884 rattus norv

ALIGNMENTS

RESULT 1

ID	AC	PRELIMINARY	PRT	690 AA.
08TAH3	08TAH3			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Similar to guanylate cyclase 1, soluble, alpha 3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-BRAIN;			
RA	Strausberg R.;			
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: BC028384; AAH28384.1; -			
SO	SEQUENCE 690 AA; 77470 MW; D9B98FC0E4DB927 CRC64;			

Query Match	Best Local Similarity	99.8%	Score 3587	DB 4	Length 690
Matches	688	Conservative	2	Mismatches	0
Indels	0	Gaps	0		
QY	1	MFTKRLDKLTGECPSFLAPQGVNSESSEKATVPICODIPEKNIDSLP	60		
DB	1	MFTKRLDKLTGECPSFLAPQGVNSESSEKATVPICODIPEKNIDSLP	60		
QY	61	ORKTSRYVLTHTLAESICKLIPPEERLVALQRTIAKHKIKESKRSEREFEXFIE	120		
DB	61	ORKTSRYVLTHTLAESICKLIPPEERLVALQRTIAKHKIKESKRSEREFEXFIE	120		
QY	121	QAAVAGVPEVVKESIGEVEFKICYEDENILGVGGLKDFLNSFTLLKQSSHQEAG	180		
DB	121	QAAVAGVPEVVKESIGEVEFKICYEDENILGVGGLKDFLNSFTLLKQSSHQEAG	180		
QY	181	KRRRLDASTLCDKRDDPLHYVFFPKRTSLILGIIKAAHVLVEVEVSLMPDCE	240		
DB	181	KRRRLDASTLCDKRDDPLHYVFFPKRTSLILGIIKAAHVLVEVEVSLMPDCE	240		

QY 241 HNDSEFVNOPIYLLYSVHMKSTKPSLSPKPOSSIVIPISLCKTEPFHFMEDKMTIQQ 300  
 DB 241 HNDSEFVNOPIYLLYSVHMKSTKPSLSPKPOSSIVIPISLCKTEPFHFMEDKMTIQQ 300  
 QY 301 FGNGIRRLNRRDFOGKPNFEFEFELLTKINOTFGITMTLMNQVAVVRMDNSYKKS 360  
 DB 301 FGNGIRRLNRRDFOGKPNFEFEFELLTKINOTFGITMTLMNQVAVVRMDNSYKKS 360  
 QY 361 SRVMDLKGOMIYVESATILFLGSPCVDRLDEFTGRLGLSLDIPINHALRDVYLIGEQR 420  
 DB 361 SRVMDLKGOMIYVESATILFLGSPCVDRLDEFTGRLGLSLDIPINHALRDVYLIGEQR 420  
 QY 421 AODGLKRRGLKATLEQAHQALEEKKKTVDLLCSIFPCVAAQOLMOGVVQAKKESNV 480  
 DB 421 AODGLKRRGLKATLEQAHQALEEKKKTVDLLCSIFPCVAAQOLMOGVVQAKKESNV 480  
 QY 481 TMLFSIYVFTALSCGSPLOYITMTLNALYTRFDQCGELDYKVTETIGDAYCVAGGLK 540  
 DB 481 TMLFSIYVFTALSCGSPLOYITMTLNALYTRFDQCGELDYKVTETIGDAYCVAGGLK 540  
 QY 541 ESDTHAVQIALMALKMMELSDVMSPHGEPKIMRIGLHSGSVAGVGVKMPRYCLFGNN 600  
 DB 541 ESDTHAVQIALMALKMMELSDVMSPHGEPKIMRIGLHSGSVAGVGVKMPRYCLFGNN 600  
 QY 601 VTLANKFESCVPRKINVSPTTYRLKDCPGFVTPRSREELPPNPFSEIPGICHFLDAY 660  
 DB 601 VTLANKFESCVPRKINVSPTTYRLKDCPGFVTPRSREELPPNPFSEIPGICHFLDAY 660  
 QY 661 OOGTNSKPCFOKKDVEDGNANFLGKASGID 690  
 DB 661 OOGTNSKPCFOKKDVEDGNANFLGKASGID 690  
 RESULT 2  
 Q9DBQ3 PRELIMINARY; PRT; 691 AA.  
 ID 09DBQ3  
 AC 09DBQ3  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE 1200016007RLK protein.  
 GN GUCY1A3 OR 1200016007RLK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=LUNG;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana H.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glasl C., King B., Kochia H.,  
 RA Kuehl P., Lewis S., Matsuo T., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszynski A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK004815; BAB23586.1;  
 DR HSP: P19687; IAWN.

DR MGD: MGI:1926562; Gucy1a3.  
 DR InterPro: IPR001054; G\_cyclase.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART: SM00044; CYCC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS01253; GUANYLATE\_CYCLASES\_2; 1.  
 KW Lyase.  
 SQ SEQUENCE 691 AA; 77587 MW; F8BDDDC3CB0BBD7F CRC64;  
 Query Match 90.9%; Score 3266.5; DB 11; Length 691;  
 Best Local Similarity 89.7%; Pred. No. 2,1e-247;  
 Matches 620; Conservative 36; Mismatches 34; Indels 1; Gaps 1;  
 QY 1 MCTRLKDKLTGCEPFLSLAPGVNPNSESEAGSSSEKATVPICODIPENKIQESLP 60  
 DB 1 MCTRLKDKLTGCEPFLSLAPGVNPNSESEAGSSSEKATVPICODIPENKIQESLP 60  
 QY 61 ORKTSRSVYVLTALAESICKLIPFEFELANALORTLAKHRTKRSKSLERDEPKITAE 120  
 DB 61 ORKTSRSVYVLTALAESICKLIPFEFELANALORTLAKHRTKRSKSLERDEPKITAE 120  
 QY 121 QAVAGVPEVYIKESIGEEVEKICEEDENILGVVGTLDLDFLNSFSTLTKQSSHCQAG 180  
 DB 121 QAVAGVPEVYIKESIGEEVEKICEEDENILGVVGTLDLDFLNSFSTLTKQSSHCQAG 180  
 QY 181 EALAGAVPEALKDSIGELFKICEDEHILGVVGTLDLDFLNSFSTLTKQSSHCQAE 180  
 DB 181 EALAGAVPEALKDSIGELFKICEDEHILGVVGTLDLDFLNSFSTLTKQSSHCQAE 180  
 QY 181 KRGLLEDAIICLKDDEFLHVIYFFPKRTSLIPGIIKAAHVIYETEVEVSLMPPCF 240  
 DB 181 KRGLLEDAIICLKDDEFLHVIYFFPKRTSLIPGIIKAAHVIYETEVEVSLMPPCF 240  
 QY 241 HNDSEFVNOPIYLLYSVHMKSTKPSLSPKPOSSIVIPISLCKTEPFHFMEDKMTIQQ 300  
 DB 241 HNDSEFVNOPIYLLYSVHMKSTKPSLSPKPOSSIVIPISLCKTEPFHFMEDKMTIQQ 300  
 QY 301 FGNGIRRLNRRDFOGKPNFEFEFELLTKINOTFGITMTLMNQVAVVRMDNSYKKS 360  
 DB 301 FGNGIRRLNRRDFOGKPNFEFEFELLTKINOTFGITMTLMNQVAVVRMDNSYKKS 360  
 QY 361 SRVMDLKGOMIYVESATILFLGSPCVDRLDEFTGRLGLSLDIPINHALRDVYLIGEQR 420  
 DB 361 SRVMDLKGOMIYVESATILFLGSPCVDRLDEFTGRLGLSLDIPINHALRDVYLIGEQR 420  
 QY 421 AODGLKRRGLKATLEQAHQALEEKKKTVDLLCSIFPCVAAQOLMOGVVQAKKESNV 480  
 DB 421 AODGLKRRGLKATLEQAHQALEEKKKTVDLLCSIFPCVAAQOLMOGVVQAKKESNV 480  
 QY 481 TMLFSIYVFTALSCGSPLOYITMTLNALYTRFDQCGELDYKVTETIGDAYCVAGGLK 540  
 DB 481 TMLFSIYVFTALSCGSPLOYITMTLNALYTRFDQCGELDYKVTETIGDAYCVAGGLK 540  
 QY 541 ESDTHAVQIALMALKMMELSDVMSPHGEPKIMRIGLHSGSVAGVGVKMPRYCLFGNN 600  
 DB 541 ESDTHAVQIALMALKMMELSDVMSPHGEPKIMRIGLHSGSVAGVGVKMPRYCLFGNN 600  
 QY 601 VTLANKFESCVPRKINVSPTTYRLKDCPGFVTPRSREELPPNPFSEIPGICHFLDAY 660  
 DB 601 VTLANKFESCVPRKINVSPTTYRLKDCPGFVTPRSREELPPNPFSEIPGICHFLDAY 660  
 QY 661 OOGTNSKPCFOKKDVEDGNANFLGKASGID 690  
 DB 661 HHOGPNSKPFQDQDVEDGNANFLGKASGID 691  
 RESULT 3  
 Q9ERL9 PRELIMINARY; PRT; 691 AA.  
 ID 09ERL9  
 AC 09ERL9  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Soluble guanylyl cyclase alpha 1 subunit.  
 GN GUCY1A3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BAIB/C;  
 RX MEDLINE=20461464; PubMed-10984516;  
 RA Sharlita I.G., Krumenacker J.S., Martin E., Murad F.;  
 RT "Genomic organization of alpha 1 and beta 1 subunits of the mammalian  
 RT soluble guanylyl cyclase genes";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:10878-10883(2000).  
 DR EMBL: AF297082; AAF17446.1;  
 DR HSSP: P19687; 1AMN.  
 DR MGI: MGI:1926562; Gucyl1a3.  
 DR InterPro: IPR001054; G\_cyclase.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART: SM00044; CYC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 DR Lyase.  
 KW SEQUENCE 691 AA; 77573 MW; F9F40C3CB0BD7F CRC64;  
 SQ

Query Match 90.9%; Score 3265.5; DB 11; Length 691;  
 Best Local Similarity 89.6%; Pred. No. 2.5e-247;  
 Matches 619; Conservative 37; Mismatches 34; Indels 1; Gaps 1;

OY 1 MCTRLNDKITGECFSLAPGVNPSSEBAAGSSSEKATVPCODIPERKNIQESIP 60  
 DB 1 MCTRLNDKITGECFSLAPGVNPSSEBAAGSSSEKATVPCODIPERKNIQESIP 60  
 OY 61 ORKTSRSNRYLHTLAESICKLIPFEERLNVALORTLAKHKIKESKSLERDEKTTAE 120  
 DB 61 ORKTSRSNRYLHTLAESICKLIPFEERLNVALORTLAKHKIKESKSLERDEKTTAE 120  
 OY 121 QAVAGVVEVEYKESLGEVEFKICEEDENILGVVGGLTKDFLNSFSTLLKSSHCQEG 180  
 DB 121 QAVAGVVEVEYKESLGEVEFKICEEDENILGVVGGLTKDFLNSFSTLLKSSHCQEG 180  
 OY 121 EAIAGAVVEALKDSBELFKICEEDENILGVVGGLTKDFLNSFSTLLKSSHCQEG 180  
 DB 121 EAIAGAVVEALKDSBELFKICEEDENILGVVGGLTKDFLNSFSTLLKSSHCQEG 180  
 OY 181 KRGLEDASICLDKEDDFLHYFFPKRTSLIPGIIKAAAHVLYETEVEVSLMPCE 240  
 DB 181 KRGLEDASICLDKEDDFLHYFFPKRTSLIPGIIKAAAHVLYETEVEVSLMPCE 240  
 OY 181 RGRLEDASICLDKEDDFLHYFFPKRTSLIPGIIKAAAHVLYETEVEVSLMPCE 240  
 DB 181 RGRLEDASICLDKEDDFLHYFFPKRTSLIPGIIKAAAHVLYETEVEVSLMPCE 240  
 OY 241 HNDSEFNOPYLLYVHMKSTKPSLSPKPOSSLVITPSLFCRPFPHFMDKMTLITQ 300  
 DB 241 HNDSEFNOPYLLYVHMKSTKPSLSPKPOSSLVITPSLFCRPFPHFMDKMTLITQ 300  
 OY 241 RSDCEFNOPYLLYVHMKSTKPSLSPKPOSSLVITPSLFCRPFPHFMDKMTLITQ 300  
 DB 241 RSDCEFNOPYLLYVHMKSTKPSLSPKPOSSLVITPSLFCRPFPHFMDKMTLITQ 300  
 OY 301 FGNGIRLRANRDFOGKPNFEFEELPKINQTFSGIMTMNMOFVVRVRRMDSVKS 360  
 DB 301 FGNGIRLRANRDFOGKPNFEFEELPKINQTFSGIMTMNMOFVVRVRRMDSVKS 360  
 OY 301 LGNGIRLRANRDFOGKPNFEFEELPKINQTFSGIMTMNMOFVVRVRRMDSVKS 360  
 DB 301 LGNGIRLRANRDFOGKPNFEFEELPKINQTFSGIMTMNMOFVVRVRRMDSVKS 360  
 OY 361 SRVMDLKGMITYVESSAILEFGSPCVRLDEFTGRGLYLSDIPIHNLRDVYLIGEQR 420  
 DB 361 SRVMDLKGMITYVESSAILEFGSPCVRLDEFTGRGLYLSDIPIHNLRDVYLIGEQR 420  
 OY 361 SRVMDLKGMITYVESSAILEFGSPCVRLDEFTGRGLYLSDIPIHNLRDVYLIGEQR 420  
 DB 361 SRVMDLKGMITYVESSAILEFGSPCVRLDEFTGRGLYLSDIPIHNLRDVYLIGEQR 420  
 OY 421 AODGKRLKGLKATLEAHQALBEKKRYDILCSITPCVAAOOLMOGOVAAKFSNV 480  
 DB 421 AODGKRLKGLKATLEAHQALBEKKRYDILCSITPCVAAOOLMOGOVAAKFSNV 480  
 OY 421 AODGKRLKGLKATLEAHQALBEKKRYDILCSITPCVAAOOLMOGOVAAKFSNV 480  
 DB 421 AODGKRLKGLKATLEAHQALBEKKRYDILCSITPCVAAOOLMOGOVAAKFSNV 480  
 OY 481 TMLFSDIYGFALICGQSPLOYITMLNLYTRFDQCGELDYKKEITIGDAYCVAGGLHR 540  
 DB 481 TMLFSDIYGFALICGQSPLOYITMLNLYTRFDQCGELDYKKEITIGDAYCVAGGLHR 540  
 OY 481 TMLFSDIYGFALICGQSPLOYITMLNLYTRFDQCGELDYKKEITIGDAYCVAGGLHR 540  
 DB 481 TMLFSDIYGFALICGQSPLOYITMLNLYTRFDQCGELDYKKEITIGDAYCVAGGLHR 540  
 OY 541 ESDTHAVOIALMALKMLSDSEVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 DB 541 ESDTHAVOIALMALKMLSDSEVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 OY 541 ESDTHAVOIALMALKMLSDSEVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 DB 541 ESDTHAVOIALMALKMLSDSEVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 OY 601 VTLANKFESSVPRKINSPITTYRLKDCPGVFTPRSRRELPRNPSPSPICHTFLAY 660  
 DB 601 VTLANKFESSVPRKINSPITTYRLKDCPGVFTPRSRRELPRNPSPSPICHTFLAY 660  
 OY 661 -OOGTNSKPCFOKDKVEDGNANFLGKASGID 690  
 DB 661 HHGGRNSKPCFOKDKVEDGNANFLGKASGID 691

RESULT 4  
 O9NNW8 PRELIMINARY; PRT; 564 AA.  
 AC O9NNW8  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE Soluble guanylate cyclase large subunit (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE OF 1-0 FROM N.A.  
 RX MEDLINE=20183694; PubMed-10717483;  
 RA Zhou Y.H., Zheng J.B., Gu X., Li W.H., Saunders G.F.;  
 RT "A novel Pax-6 binding site in rodent B1 repetitive elements:  
 RT coevolution between developmental regulation and repeated elements";  
 RL Gene 245:319-328(2000).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA Zhou Y.H.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF233750; AAF64043.1;  
 DR EMBL: AF233746; AAF64043.1; JOINED.  
 DR EMBL: AF233747; AAF64043.1; JOINED.  
 DR EMBL: AF233748; AAF64043.1; JOINED.  
 DR EMBL: AF233749; AAF64043.1; JOINED.  
 DR HSSP: P19687; 1AMN.  
 DR InterPro: IPR001054; G\_cyclase.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART: SM00044; CYC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 DR Lyase.  
 KW NON-TER  
 FT SEQUENCE 564 AA; 63406 MW; 7056E249A75F98D5 CRC64;  
 SQ

Query Match 82.2%; Score 2954; DB 4; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-223;  
 Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 VVEVYKESLGEVEFKICEEDENILGVVGGLTKDFLNSFSTLLKSSHCQEGRGRLE 186  
 DB 1 VVEVYKESLGEVEFKICEEDENILGVVGGLTKDFLNSFSTLLKSSHCQEGRGRLE 186  
 OY 187 DASICLDKEDDFLHYFFPKRTSLIPGIIKAAAHVLYETEVEVSLMPCEFHNDSE 246  
 DB 187 DASICLDKEDDFLHYFFPKRTSLIPGIIKAAAHVLYETEVEVSLMPCEFHNDSE 246  
 OY 247 FVNOPYLLYVHMKSTKPSLSPKPOSSLVITPSLFCRPFPHFMDKMTLITQNGIR 306  
 DB 247 FVNOPYLLYVHMKSTKPSLSPKPOSSLVITPSLFCRPFPHFMDKMTLITQNGIR 306  
 OY 307 RLNRNDRDFOGKPNFEFEELPKINQTFSGIMTMNMOFVVRVRRMDSVKSRYMDL 366  
 DB 307 RLNRNDRDFOGKPNFEFEELPKINQTFSGIMTMNMOFVVRVRRMDSVKSRYMDL 366  
 OY 367 KCOMITYVESSAILEFGSPCVRLDEFTGRGLYLSDIPIHNLRDVYLIGEQRADGILK 426  
 DB 367 KCOMITYVESSAILEFGSPCVRLDEFTGRGLYLSDIPIHNLRDVYLIGEQRADGILK 426  
 OY 427 KRLGKRLKATLEAHQALBEKKRYDILCSITPCVAAOOLMOGOVAAKFSNVTMLFS 486  
 DB 427 KRLGKRLKATLEAHQALBEKKRYDILCSITPCVAAOOLMOGOVAAKFSNVTMLFS 486  
 OY 487 IYGFALICGQSPLOYITMLNLYTRFDQCGELDYKKEITIGDAYCVAGGLHKSSTHA 546  
 DB 487 IYGFALICGQSPLOYITMLNLYTRFDQCGELDYKKEITIGDAYCVAGGLHKSSTHA 546  
 OY 547 VOIALMALKMLSDSEVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNNVTLANK 606  
 DB 547 VOIALMALKMLSDSEVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNNVTLANK 606

DB 421 VOIATAMALAMELSDVEMSPHGEPIKRIKRIHSGSVAGVYVAKMPYICLFGNNVTLANK 480

QY 607 FESGCVPRKINVSPTTYRLKDCPGFVFTRPRELPNPPSEIPGICHELDAVQOQTS 666

DB 481 FESGCVPRKINVSPTTYRLKDCPGFVFTRPRELPNPPSEIPGICHELDAVQOQTS 540

QY 667 KPCFOKKDVEDGNANFLGKASGID 690

DB 541 KPCFOKKDVEDGNANFLGKASGID 564

RESULT 5

Q9PW12 PRELIMINARY; PRT; 678 AA.

ID Q9PW12

AC Q9PW12; 01-MAY-2000 (TREMREL. 13, Created)

DT 01-MAY-2000 (TREMREL. 13, Last sequence update)

DT 01-MAR-2002 (TREMREL. 20, Last annotation update)

DE Soluble guanylyl cyclase alpha subunit.

OS Oryzias latipes (Medaka fish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; OC Belontiiformes; Adriantichthyidae; Oryziinae; Oryzias.

OX NCBI\_TaxID=8090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-99303623; PubMed-10373466;

RA Mikami T., Kusakabe T., Suzuki N.;

RT "Random organization of medaka fish soluble guanylyl cyclase alpha and beta subunit genes. Implications for coordinated transcription of two subunit genes."

RT J. Biol. Chem. 274:18567-18573(1999).

RL EMBL; AB022280; BAA76690.1; -.

DR HSSP; P19687; IAWN.

DR InterPro; IPR001054; G\_cyclase.

DR Pfam; PF00211; guanylate\_cyc; 1.

DR SMART; SM00044; CYCG; 1.

DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.

DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 1.

KW Lyase.

SQ SEQUENCE 678 AA; 75166 MW; 66623709CBE68A5C CRC64;

Query Match 56.18; Score 2014.5; DB 13; Length 678;

Best Local Similarity 59.96; Pred. No. 2.8e-149;

Matches 403; Conservative 99; Mismatches 152; Indels 19; Gaps 8;

QY 1 MECTKLKDKITGECPSFLAPGQVPNESEEAAGSSCAKATVPICODIPKNIQESLP 60

DB 1 MFCALKEIKLISGECPPSSAKN---NELGDFEVRSSDAAD-SLPISEVHGK-IRENVP 55

QY 61 QRTSRSRYTLHTLAESICKLIPFERLVALQR--TLAKRIKESRSL-----ERE 112

DB 56 ROKASRAKVMHTLGDISIRKLACPGFORLHDALQRMVTAADHSKSESTSLHCHYQSCS 115

QY 113 DEKRTIAEQ---AVAAGVVEYIKSLGEVFKICEEDENILGVGGLKPLNFSF 168

DB 116 DKRHLVDMMKSYSTKTAIPMEALKITALGEIFSMCEEGHILRYVGGALHDFLNSFN 175

QY 169 LKQSSHCQAGRGRLSDASILCLDEDFLHVYFFPKRTSLILPGIKAANVLYE 228

DB 176 LKQSSMLPSPREDCCNEBSVCLDKDPDLLVYFENPSPTEMPFGVYKAAACLLYN 235

QY 229 TEVEVSLMPPCFNDCSEFNQPYLLYSVMKSTKP-SLSPSKPQSSLVIPSLFCKTFP 287

DB 236 TVYDVVLKPLTSDISLQSSPQSLTYTVVAKAKTISPRLRATSGTLPSTLSFTFP 295

QY 288 FHMFMKDMITLQFNGIRLNMRRD-FQCKPNEFEFFELTPKINTEFGIMTMNMOF 346

DB 296 FHLLDQDLVLYIGGLRRRLTRKGLRSPFQEFISVSPQIKTFGGLITMTNTOF 355

QY 347 VVVRWMDNSVKKSRMDLKGOMIYVESATILFGSPVDRLDEDTGGLVLSIDPIH 406

DB 356 IIRIKGVFTDNTGKPMIDIKGOMIYVESNAIILFLGSPCVDLKEELTGGLYLSIDPIH 415

QY 407 NALRDVYLIGEQAQADGLKRRKRLKATLEQAQHAOLEEKKRTVDLCSIFPCEVAQOL 466

DB 416 NALRDVYLIGEQAQADGLKRRKRLKATLEQAQHAOLEEKKRTVDLCSIFPCEVAQOL 475

QY 467 WQGVQAKRFNSVMTLSDIVGFTALCSQSPLOVITMLALYTRPDQCGELDYKYVE 526

DB 476 WQGVQAKRFNSVMTLSDIVGFTALCSQSPLOVITMLALYTRPDQCGELDYKYVE 535

QY 527 TIGDAVAGGLKRESPTHAVOIATAMALAMELSDVEMSPHGEPIKRIKRIHSGSVAGV 586

DB 536 TIGDAVAGGLKRESPTHAVOIATAMALAMELSDVEMSPHGEPIKRIKRIHSGSVAGV 595

QY 587 VGVKMPRYCLFGNNVTLANKFESGCVPRKINVSPTTYRLKDCPGFVFTRPRELPNPP 646

DB 596 VGVKMPRYCLFGNNVTLANKFESGCVPRKINVSPTTYRLKDCPGFVFTRPRELPNPP 655

QY 647 PSEIPGICHELDA 659

DB 656 PEDIPGVCYFLES 668

RESULT 6

P79997 PRELIMINARY; PRT; 678 AA.

ID P79997

AC P79997; 01-MAY-1997 (TREMREL. 03, Created)

DT 01-MAY-1997 (TREMREL. 03, Last sequence update)

DT 01-MAR-2002 (TREMREL. 20, Last annotation update)

DE Soluble guanylyl cyclase alpha subunit (EC 4.6.1.2).

OS Oryzias latipes (Medaka fish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; OC Belontiiformes; Adriantichthyidae; Oryziinae; Oryzias.

OX NCBI\_TaxID=8090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-98237571; PubMed-9578459;

RA Mikami T., Kusakabe T., Suzuki N.;

RT "Molecular cloning of cDNAs and expression of mRNAs encoding alpha and beta subunits of soluble guanylyl cyclase from medaka fish Oryzias latipes."

RT Eur. J. Biochem. 253:42-48(1998).

RL EMBL; AB000849; BAA19198.1; -.

DR HSSP; P19687; IAWN.

DR InterPro; IPR001054; G\_cyclase.

DR Pfam; PF00211; guanylate\_cyc; 1.

DR SMART; SM00044; CYCG; 1.

DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.

DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 1.

KW Lyase.

SQ SEQUENCE 678 AA; 75133 MW; 3A6141D2A6475D40 CRC64;

Query Match 55.98; Score 2009.5; DB 13; Length 678;

Best Local Similarity 59.68; Pred. No. 7e-149;

Matches 401; Conservative 101; Mismatches 152; Indels 19; Gaps 8;

QY 1 MECTKLKDKITGECPSFLAPGQVPNESEEAAGSSCAKATVPICODIPKNIQESLP 60

DB 1 MFCALKEIKLISGECPPSSAKN---NELGDFEVRSSDAAD-SLPISEVHGK-IRENVP 55

QY 61 QRTSRSRYTLHTLAESICKLIPFERLVALQR--TLAKRIKESRSL-----ERE 112

DB 56 ROKASRAKVMHTLGDISIRKLACPGFORLHDALQRMVTAADHSKSESTSLHCHYQSCS 115

QY 113 DEKRTIAEQ---AVAAGVVEYIKSLGEVFKICEEDENILGVGGLKPLNFSF 168

DB 116 DKRHLVDMMKSYSTKTAIPMEALKITALGEIFSMCEEGHILRYVGGALHDFLNSFN 175

QY 169 LKQSSHCQAGRGRLSDASILCLDEDFLHVYFFPKRTSLILPGIKAANVLYE 228

```

Db 176 LIAKSSMLPSPDRDCVNEPSVLCIDDPCLLYTFENPSPTEMEFPGVIAKAAACLIYN 235
QY 229 TEVEVSILMPCFHNDCESEFVNOPYLLYSVHKSTKP-SLSPKPOSILVPTSLFCKTP 287
Db 236 TTVVLDVLPATKDSILQSSPQSLTYTVVAKAKTLSPSLRATSNAGTLSTLSTFP 295
QY 288 FHFEPKDMTILQFGNGIRLRMRD-FQKPNFEYFELLTPKINOTSGIMTMNM 346
Db 296 FHLIDDLVLYVOGHGRKRLTRKDLRSPTEQHEHSIVSPQIKCTGCLTMANTQF 355
QY 347 VVRVRMDNSVKKSRVMDLKGOMIYVESAILFLGSPCYDRLEDFTGRGLYSDPIH 406
Db 356 IIRKHEVETDNTGKPMDLKGOMIYVESNAIIFLGSPCYDKLEELTGRGLYSDPIH 415
QY 407 NALRDVYLIGEOARADGLKRLKATLEQAHAQALEEKKRTVDLLCSIFPEVAQOL 466
Db 416 NALRDVYLIGEOARADGLKRLKAKALENAQALEEKKRTVDLLCSIFPEVAQOL 475
QY 467 WOGVOVAKKFSNTMLFSDIVGFTALCSQCSPLQVITMLALYTRDQCGELDYKVE 526
Db 476 WOGTVOAKKFNQYTMFSDIVGFTAVCSLCTPMQVITMLNELYTKFQYCGELDYKVE 535
QY 527 TIGDAYCVAGGLHESDTHAVOIALMALKMMELSDVMSPHGEPITKMRIGLHSGSVFAGV 586
Db 536 TIGDAYCVAGGLHESDTHAVOIALMALKMMELSDVMSPHGEPITKMRIGLHSGSVFAGV 595
QY 587 VGVKMPRYCLFGNNVTLANKEFSCSVPRKINVSPTTYRLKDCPGVFTPRSRRELPP 646
Db 596 VGVKMPRYCLFGNNVTLANKEFSCSVPRKINVSPTTYRLKDCPGVFTPRSRRELPP 655
QY 647 PSEIPIGICHLDA 659
Db 656 PEDIPGVCYFLES 668

```

## RESULT 7

```

QY 090VVS PRELIMINARY: PRT: 675 AA.
AC 090VVS;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Soluble guanylyl cyclase alpha subunit.
GN FRCS-ALPHA1.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
CX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Morinaga C., Yamamoto T., Moriya Y., Suzuki N.;
RT "Identification of tandem organization of soluble guanylyl cyclase
RT alpha and beta subunit genes in the Japanese pufferfish (Fugu
RT rubripes) genome."
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062171; BAB60907.1; -
DR EMBL; AB062169; BAB60905.1; -
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; UNKNOWN_1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
SQ SEQUENCE 675 AA; 75498 MW; E71A283DC0369601 CRC64;

```

```

Query Match 55.8%; Score 2005; DB 13; Length 675;
Best Local Similarity 60.1%; Pred. No. 1.6e-148;
Matches 406; Conservative 96; Mismatches 147; Indels 26; Gaps 9;

```

```

QY 1 MFCRLDKLTGTCGPFSLA-PCQVNPSESEAGSESCATVPTICODIPKKNIQSL 59
Db 1 MFCRLKELKISGCPFTSSAKTNOL-EHSEQLSKD---VAHLPTSKDV-RRTTGNA 53

```

```

QY 60 PQRSTSRVYLHTLAESICLIPPEPERLNVALQRTLAKK-----IKESRK 107
Db 54 PHKSSRAKVMILHTLGESIRKLACPEPERLHTLRMRRLSDPSREPAFVCCTEYQRC 113
QY 108 SLEREDFKTIAEOAVAGVVEYIKESLGEVYKICYEEDENILGVGTLKDFLNS 167
Db 114 SDDEHFFVKMMSMRRAIMDRLVALGELFMCCYEEDRHILRVGGALHFLNSFN 173
QY 168 TLLKSSQCRAGKRGRLDEASIIICLDEKDEFLHYFFPRRTSLILGIIKAAHYLY 227
Db 174 VLLKQSLTHIQDRKDCVNEPSVLCIDDLGLLYTFENPHTTELEFPVIAKAAARLY 233
QY 228 TEVEVSILMPCFHNDCESEFVNOPYLLYSVHKSTKPSLSPS--KPSLSVLPSTLCKT 285
Db 224 HTVVDVLDLPATKDSILQSSPQSLTYTVVAKAK-TLSPLRATSNAGTLSTLSTFP 292
QY 286 FHFEPKDMTILQFGNGIRLRMRDPOKRP-NFEYFELLTPKINOTSGIMTMNM 344
Db 293 FHLIDDLVLYVOGHGRKRLTRKDLRSPTEQHEHSIVSPQIKCTGCLTMANTQF 352
QY 345 QEVVRMRDMSVKKSRVMDLKGOMIYVESAILFLGSPCYDRLEDFTGRGLYSDPI 404
Db 353 QFTIRKGVSTA--DNTLMDLKGOMIYVESDAIIFLGSPCYDKLEELTGRGLYSDPI 410
QY 405 IHNALRDVYLIGEOARADGLKRLKATLEQAHAQALEEKKRTVDLLCSIFPEVAQ 464
Db 411 IHNALRDVYLIGEOARADGLKRLKAKALENAQALEEKKRTVDLLCSIFPEVAQ 470
QY 465 QLMGQVOVAKKFSNTMLFSDIVGFTALCSQCSPLQVITMLALYTRDQCGELDYK 524
Db 471 QLMGQVOVAKKFSNTMLFSDIVGFTAVCSLCTPMQVITMLNELYTKFQYCGELDYK 530
QY 525 VETIGDAYCVAGGLHESDTHAVOIALMALKMMELSDVMSPHGEPITKMRIGLHSGSV 584
Db 531 VETIGDAYCVAGGLHESDTHAVOIALMALKMMELSDVMSPHGEPITKMRIGLHSGSV 590
QY 585 GVGKMPRYCLFGNNVTLANKEFSCSVPRKINVSPTTYRLKDCPGVFTPRSRRELPP 644
Db 591 GVGKMPRYCLFGNNVTLANKEFSCSVPRKINVSPTTYRLKDCPGVFTPRSRRELPP 650
QY 645 NPEIPIGICHLDA 659
Db 651 NPEIPIGICHLDA 665

```

## RESULT 8

```

QY 077105 PRELIMINARY: PRT: 699 AA.
AC 077105;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Soluble guanylyl cyclase alpha-1 subunit.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
CX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-98409773; PubMed-9736646;
RA Nighorn A., Gibson N.J., Rivers D.M., Hildebrand J.G., Morton D.B.;
RT "The nitric oxide-cGMP pathway may mediate communication between
RT sensory afferents and projection neurons in the antennal lobe of
RT manduca sexta."
RL J. Neurosci. 18;7244-7255(1998).
DR EMBL; AF062750; AAC61263.1; -
DR HSSP; P19687; IAWN.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYC; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.

```



KM Lyase. 699 AA; 78490 MW; 989F8067ABAF8465 CRC64;  
 SQ SEQUENCE

Query Match 30.5%; Score 1097; DB 5; Length 699;  
 Best Local Similarity 37.6%; Pred. No. 2,5e-77;  
 Matches 266; Conservative 130; Mismatches 262; Indels 50; Gaps 19;

15 CFF-----SLAPGVNPNSSBAAGSSSCATVPICOD-----IPKNIQESLPQRK 63  
 DB 3 CFFRRASOHOFGANGSSAPAKKPEFRSRSSVHLTGPEEDGRNLTLLHKMSAL-QTL 61

64 TSRSRYLHTLASSICKLIPPEFERLNVALQRLANHKIKESHRSLEDEFEKTIAQAV 123  
 DB 62 TAPSNECLHAAYTSLRKNSDTHKKN-----CLRLPDVYKCNNAVYLOEITDAVR 114

124 AA-GVPEVVEIKESIGEVKICYEED---ENILVGVTGLKDFLNSFTLLKSSHQEA 179  
 DB 115 ATDSVNTKDFPMATIGYLLITANSHNCRLEAKFCGTNLTLETLTIDS-VHYVLDHDT 173

180 G-KRGRLD-DASTLCDEDEFLVHYFFPKRTSLILPGIKAHAHVETEVEVSLMP 237  
 DB 174 PLKDETFMEYEAANVCTTSEGGKIQHLHTESEPAVAVLSLKAIAKRLDTOTDIRLR- 232

238 PCFHNDCSEFVNPVLLYSY--HMKSTR-----PLSPSKPQSSIVIPSTLCKPT 286  
 DB 233 -STNDPRRF--RYELINAVPLHOKSKESCELVNENASAVSTKVTDLKIGVASCKAP 288

287 PEHFMDKMTILQFNGIRLNNRDFQKPNFEFEFLTPK-INQFSGIMTLMNQ 345  
 DB 289 PMHFTDKRLVELQAGFMRLLFGTHLHAGSLGYFRLNRGVPLDRREILKKNTP 348

346 FVYRVRRMNSVAKSSRWDLKCOMIYESSAILFSGSCVRLDFGRTGLYSLDIP 405  
 DB 349 FMECLMPSTALAE--LEIKQWFCESDLSLGFDFGLRGLRGFISDIP 406

406 HNALRDVYLIGEQARADGKRLKATLEQAHOLEBEKKKYVDLCSIPCEFAQ 465  
 DB 407 HDHTRVYILVGEARADGKRLKATLEQAHOLEBEKKKYVDLCSIPCEFAQ 466

466 LMGOVVOAKKFSNVTMLFSDIVGTFATCSQSPLOYITMLNLYTRFDQCGELDYKY 525  
 DB 467 LMTGERIEAKSHDVTMLFSDIVGTFATCSQSPLOYITMLNLYTRFDQCGELDYKY 526

526 ETIGDVAACGGLKRESDFHVAOIALMALKMLSDSEVNSPHREPIKMRIGLSSGVAC 585  
 DB 527 ETIGDVAACGGLKRESDFHVAOIALMALKMLSDSEVNSPHREPIKMRIGLSSGVAC 586

586 VVGKMPRYCLFEGNNVTLANKFESSVPRKINSPTTYRLKDCPGFVTPRSREELPN 645  
 DB 587 VVGKMPRYCLFEGNNVTLANKFESSVPRKINSPTTYRLKDCPGFVTPRSREELPN 646

646 FPESEIPGICHTLDAV-QQGTN-SKPCFQKKDVEDGNANF-LGRKASGID 690  
 DB 647 FPESEIPGICHTLDAV-QQGTN-SKPCFQKKDVEDGNANF-LGRKASGID 692

RESULT 9  
 Q24085 PRELIMINARY; PRT; 676 AA.  
 ID Q24085  
 AC Q24085;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE GYCALPHA99B protein.  
 GN GYC-ALPHA-99B OR GYCALPHA99B OR DGCA1 OR CG1912.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 [1]  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;

RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Cealiner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Blazet R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,  
 RA Artil J.F., Agdayani A., An H.-G., Andrews-Planckoch C., Baldwin D.,  
 RA Beeson R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson R.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,  
 RA Borokva D., Botchan M.R., Bouck J., Brockett P., Brothier P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA De Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin R.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spter E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Welschbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OREGON-R;  
 RX MEDLINE-95318108; PubMed-7797526;  
 RA Shah S., Hyde D.R.;  
 RT "Two Drosophila genes that encode the alpha and beta subunits of the  
 RT brain soluble guanylyl cyclase.";  
 RL J. Biol. Chem. 270:15368-15376(1995).  
 DR EMBL: AE003770; AAF56917.1;  
 DR EMBL: U27117; AAB7940.1;  
 DR HSSP: P19687; 1AAN.  
 DR Flybase: FBgn0013972; Gyc-alpha-99B.  
 DR InterPro: IPR001054; G\_cyclase.  
 DR Pfam: PF00211; guanylate\_cyc.  
 DR SMART: SM00044; CYCG; 1.  
 DR PROSITE: P550125; GUANYLATE\_CYCLASES; 2; 1.  
 SQ SEQUENCE 676 AA; 75662 MW; CE8097ELK3787F8 CRC64;

Query Match 27.6%; Score 991.5; DB 5; Length 676;  
 Best Local Similarity 37.8%; Pred. No. 4,4e-69;  
 Matches 247; Conservative 117; Mismatches 209; Indels 81; Gaps 20;

58 SLPORTSRYVLTFLASICKLIPPEFERLNVALQRLAK-----HKIK---ESR 106  
 DB 27 ALEDELSDALTLTHQMAIOLITPAPSNEDLNTAVTSVAXYRQMPNIIHKIKLDPQF 86

107 KSLREDFEKTIAE-----QAVAGVPEVVEIKESIGEVKICYEED-DENILGVYGL 159  
 DB 87 KSCANTDIADLOELILKMDASASITLV-----LGEELITCCGCTIIRAFRCGTDL 141

160 KDFLNSFS---TLKSSHCQENGKRGRLDASILCLDEDEDFLHYFFPKRTSLIP 216  
 DB 142 QEFISGLSDGVYDLK-----LQED-----VTDGFCVAGGGE-----LIFSERVIANMLL 189



QY	217	G1TAAATVETEEVSLIMPPCFINDCSEFNOPYLATSVH-----MKTSP-SLSPK	210
Db	190	GSIALATLMALKVDVNIKIEP--VEGDARY--RTLSVLADNSOTMLMGRTYSVKTI	244
QY	271	PQ-----SSVIPFSLFCKTPEPHFMEDKDMTILQFNGIRLNRDRFOGKPNF	320
Db	245	PETVGRSNSSNADSLQNMSSSPCKMEPHFIMNEQLEVLQGRGFSKLT-----KPYN	297
QY	321	EE-----YEILTPK-INQTFSSIMTMLNQFVVRARARNDSVAKSSRWDLKCOMIY	372
Db	298	ADFGCQATVTVDFKRPKGLTKMKFRIVRRTYTPFLIGLNNPQGAADFPAIGLEIKQWHA	357
QY	373	IVESSAIFLTPSPCYDRLEDFTGRLQSLSDIPRIHNAIRDVYLIGEORAAQDGLKRLGTL	432
Db	358	CPENSNILFITSPLFDGLDGLTGNLFLSPDILPHATREYILVGEARARQDGLRRMDKI	417
QY	433	KATLEQAQALBEEKKTKVDLLCSITPECEVAQOLMOGOVVAQKKSNTMTLSPDIVGFTA	492
Db	418	KNSIEEANSATVTKERKNVSLHLHIFPEAIEKMLGSSIDAKTYPDVTLIFSDIVGFTS	477
QY	493	ICSOCSPLQVTVTMLNALTATRFDOCGEILDVYKVENTIGDAVCYAGGILHKSQDPHAYOIALM	552
Db	478	ICSRATPMTVLSMLEGLKKDEDFEDFYVEITIGDAVCYAGGILHRSIYDAHKVYAWA	537
QY	553	ALKMELSDENVSPHGEPIKRIKRIGLSHSSVFAGVGVKMPRYCLFENNTYLANKEFSCSV	612
Db	538	ALKMIDACSKHITHDGEQIKRRIGLHTGTVLAGVGRKMPRYCLFCHSVTIANKEFSGSE	597
QY	613	PRKINVSFTYTRLLKDCGPFV--TPRSEELPPNPSPSITPG---ICHLDAVQ	661
Db	598	ALKINVSFTTDMWLTKHGEFEELQRPDPSPFLPKPEPN--PGTETCYPLESRF	649
RESULT 10			
Q958Q4			
ID	Q95Q4	PRELIMINARY:	PRT: 676 AA.
AC	Q95Q4:		
DT	01-DEC-2001	(TREMBLrel. 19, Created)	
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)	
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)	
DE	GH08311P.		
GN	GYC-ALPHA-99B OR GYC-ALPHA99B OR CG11912.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
XX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RA	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,		
RA	Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,		
RA	Gonzalez M., Guarin H., Li P., Liso G., Miranda A., Mungall C.J.,		
RA	Nunoo J., Pacle J., Paragas V., Park S., Phuananavong S., Wan K.,		
RA	Yu C., Lewis S.E., Rubin G.M., Celniker S.,		
RL	Submitted (Oct-2001) to the EMBL/Genbank/DBU databases.		
DR	EMBL; AY060654; AAL28202.1; -		
DR	FLYbase; FBgn0013972; GYC-alpha-99B.		
DR	InterPro; IPR001054; G_Cyclase.		
DR	Pfam; PF00211; guanylate_cyc; 1.		
DR	PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.		
SO	SEQUENCE 676 AA: 75658 MW: CB9183E1EC314138 CRC64:		

Query Match	27.4%;	Score 985.5;	DB 5;	Length 676;
Best Local Similarity	37.6%;	Pred. NO. 1.3e-68;		
Matches 246;	Conservative 117;	Mismatches 210;	Indels 81;	Gaps 20

QY 58 SLPOKRTSRKRYLHTLTAESICKLFFPERERLVALQRLAK-----HKRK-----ESR 106  
QY 27 ALEDELSDDATLTHLMAIOLLTPASNEDLNTAVTSLAKYRONPNHKKLDPQTE 86  
QY 107 KSLSEDEDEKTLAE-----QAVAGVPRPEVAKESGGEVFKCYEE-DENILGVGGTL 155

```

Db      87 KSCANDYIADIQELLKMDKDEASAEIly-----LlGEllLTCCGIIIEARRCGLTD 141
Qy      160 KDFLNFs---TllKOSshCQAGKRGLEdASIlLClDKEDdELhlyFFPPKRTSlLlP 216
Db      142 QeFlTGLDGVyVLK-----lQEBED---vTDTGfVCAGeGE---lTfTSERPvIAMllL 189
Qy      217 GIllKAANhlyETEVEvSLMPPCFlNDCSERvNOPIllLYsh-----MKTpR-SlSPSK 270
Db      190 GSlKAlTRMLyRVDVNIKElP--VEGDARRy---RYLESlLKONSQTMlNGRPlTSVKIT 244
Qy      271 PQ-----SlyVlPTSLCkTFFPHEMDKMTlLQNGNlRlRLMRRDPOGPN 320
Db      245 PElVQNSNSNSMsdLOMNSStCKKFPNHfIMNEDElEYlQGRGfSkly-----KPIY 297
Qy      321 EE-----YFElLlPK-INQTFSGlMTLMNQFvYVRMRDMSYKKSsRYMDlKGMlY 372
Db      298 AdEGCQATyFPFKRkRGlTMKFRDlVRRTYTPFlGlNPNPGAvDPAlGLElKGmYH 357
Qy      373 IYESSAlLlLGSPCVDRLedFTGRlLYsDlPlRlNALdVYLlGEQARADGlKkRlGKL 432
Db      358 CPESNSlLlTGSFPlDGlDGlCNlElfSDlPlHDATeRvLYlVEQARADGGLRRMDKl 417
Qy      433 KAlEQAOhAlEeEKkTYdLlCSlFPECVeAQOllMOGVQVAKKFSvNTMlFSlYGEfA 492
Db      418 KNSIEERANSvYlKERKKvNSlLHlLPRAIEAKlMLGSSlDAKTYPPvTlIFSDlYGFIS 477
Qy      493 lCSQCSPlOvYlTMlNALYlRFDQOGELdYVKvETlGDAYCVAGLhRESDTRAVOYAlM 552
Db      478 lCSRAPlRPNvYlSMlLBEGLYKDFEDFCDFDPDYKvETlGDAYCVASGLhRASlYOhAKYAM 537
Qy      553 AlMAMeLSDEvMSPhCEPlKARlGlHSGSVFAGvYAKMPRYClFGNNVtlANKESCSv 612
Db      538 AlMAlDACSKhIITHQEOlKMRlGlHTGfVlAGvYGRMPRYClGHGSvIYANKFESGSE 597
Qy      613 PRKlNVSPtTYRlLKDCQGEfV--lPRShEELPRvPESElTG---lOhPlDAYO 661
Db      598 AlKlNVSPtPKDMLlTHGCFEElLOPRDPSlPKFEPR--PGGfETCYlFESFR 649

```

RESULT	ID	QBR514	PRELIMINARY;	PRT;	220 AA.
AC	QBR514				
DC	QBR514				
DT	01-JUN-2002	(TEMBirel. 21, Created)			
DT	01-JUN-2002	(TEMBirel. 21, Last sequence update)			
DT	01-JUN-2002	(TEMBirel. 21, Last annotation update)			
DE	Soluble guanylate cyclase alpha2d (Fragment).				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID-10116;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SPRAGUE-DAWLEY;	TISSUE-BRAIN;			
RA	Okamoto H., Asakawa T.;				
RT	"Molecular cloning of a novel variant of the rat soluble guanylate cyclase beta2 subunit."				
RL	Submitted (FEB-2002) to the EMBL/genbank/DBJ databases.				
DR	EMBL, AB079780, BAB84824.1;	-.			
FT	NON_TER	1			
SO	SEQUENCE	220	220		
		220 AA;	24646 MW;	DBECC749FDA5E35E CRC64;	

Query Match	22.98;	Score 823;	DB 11;	Length 220;
Best Local Similarity	60.78;	Pred. No. 1.5e-56;		
Matches 162; Conservative	30;	Mismatches 27;	Indels 48;	Gaps 1.

[illegible]

DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	
----	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

[illegible]

```

OY 358 --KSSRYMDKGMITYVSSAIIIFGSPCVDRLEDTGRLYISDIPINHALDVL 415
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 287 DPHEIASLRKGMILYIPETDVVVFQCPSTNIDDLTRGLCIADIPLDHATRDVL 346
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 416 GEQARADGKLRKLGKATLEQAHQALEEKKKTVDLLCSIFPEVAQOQMOGVQAK 475
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 347 SEQPEADIKLTQNLVLTQKQFFRELEERKQKTRLLSVLPISVATELRHRRPVAR 406
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 476 KFSNVTMLSDIVGFTALCSOCSS--PLQVITMLNALTFRPD--QCCGELDYKVEI 528
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 407 RYDVITLFSGIVGFTALCSOCSS--PLQVITMLNALTFRPD--QCCGELDYKVEI 466
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 529 GDACVAGGLKESPDTHAVQIALMALKMEISDEVMSPHGPRIKRIGLHSGVAFAGV 588
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 467 GDKYVAVGVGVVYVESHQORVAVNFRALGRISAKFMANVYGPDIORIGIRGPVLAGV 525
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 589 VKMPRYCLFGNNVTLANKEFSCSVPRKINVSPTTYRL 626
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 526 HKMPRYCLFGNNVTLANKEFSCSVPRKINVSPTTYRL 563
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 14

091XJ7 PRELIMINARY: PRT: 742 AA.

```

ID 091XJ7
AC 091XJ7
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Soluble guanylyl cyclase beta 2 subunit.
GN GUCY1B2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=21391941; PubMed=11406623;
RA Koglin M., Vohse K., Budaus L., Scholz H., Behrends S.;
RT "Nitric Oxide Activates the beta 2 Subunit of Soluble Guanylyl Cyclase
  in the Absence of a Second Subunit."
RL J. Biol. Chem. 276.30737-30743(2001).
DR EMBL; AF004153; AAF6581.1; -.
DR InterPro; IPR001054; G_cyclase.
DR InterPro; IPR001230; Prenyl_site.
DR Pfam; PF00211; guanylate_cyc; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; UNKNOWN_1.
DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
SQ SEQUENCE 742 AA; 83050 MW; 41CE148B15E244E4 CRC64;

```

Query Match 21.8%; Score 784.5; DB 11; Length 742;  
 Best Local Similarity 32.6%; Pred. No. 8.5e-53;  
 Matches 194; Conservative 107; Mismatches 213; Indels 81; Gaps 10;

```

OY 116 KTAIEQAVAGVPEVETIKESIGEYFKICYED-ENIGVVGGLKDLNFESESTLLK 174
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 48 KLIOEACKVLDSMEALIKLFGEYFFKCKMSGYRMLRTLGNTLETFENDAL----- 102
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 175 HGOEAGKRGRLSDSILCLDKEDDFLHYTFPPKTTSLIIPGIIKAAHVLTEFEVYS 234
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 103 HSYTALSTQENNAFSFVEEGADGMLLHYSDRGHLCHVPGITIEAVAKDFDIDVMS 162
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 235 LMPCCFHNDGSEFV---NOPYLLSYVMKSTRP--SLSPKPOSS----- 274
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 163 IL-----DMEVEVETKKEHVFLVYQAKHROIKGAASRPOSSQADQDLQGL 216
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 275 -----LVPTSLFCKTFEPH 289
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 LRMKERYLINIPVCPGEKSHSTAVRASVLEFGKPLDFTQPVYPERLWVEEVEFCAPFPH 276
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 290 FMFQKDMITLQFGNCRILMLNRDPOGKPNPEEYFEIILPKINQIFSGIMTLMNQFVVR 349
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 277 IVFDEALRKQAGVNIQKYVGLITQ--KFADEFFSIHPQVTPNISSICKFINSQFVLK 335
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 350 VRR--WDSNKSSRYMDKGMITYVSSAIIIFGSPCVDRLEDTGRLYISDIPINHA 408
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 336 TREKMPARAKSOPRLKRGIMIMESLRCHIFMCSPPVNSLOELESKMLSDIAPDPT 395
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 409 LRDVVLIGEQARADGKLRKLGKATLEQAHQALEEKKKTVDLLCSIFPEVAQOQMO 468
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 396 TRDILLNQORLAEMELSCQLEKKKEELRVLSNHLIAETKKTETLLVAMPDEHVAHQKE 455
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 469 GOVYQAKFNSVTMLSDIVGFTALCSOCSS--PLQVITMLNALTFRPDQCCGELDYKVEI 528
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 456 GRKVAAGEFETCTILFSDVVTFTNCAACEPIQIVNMNMSYKFEIDLTSVHDYKVEI 515
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 529 GDACVAGGLKESPDTHAVQIALMALKMEISDEVMSP--HGPRIKRIGLHSGVAFAGV 587
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 516 GDATVAVGVGVVYVESHQORVAVNFRALGRISAKFMANVYGPDIORIGIRGPVLAGV 575
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 588 VKMPRYCLFGNNVTLANKEFSCSVPRKINVSPTTYRLKDCPGFVTPRSREEL 642
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 576 GDKMPRYCLFGNNVTLANKEFSCSVPRKINVSPTTYRLKDCPGFVTPRSREEL 629
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 15

054865 PRELIMINARY: PRT: 620 AA.

```

ID 054865
AC 054865;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Soluble guanylyl cyclase beta-1 subunit.
GN GUCY1B3 OR GC-S-BETA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX PubMed=10984516;
RA Gansseman Y., Brouckaert P., Fiers W.;
RT Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/10;
RX PubMed=10984516;
RA Sharita I.G., Krumenacker J.S., Martin E., Murad F.;
RT "Genomic organization of alpha 1 and beta 1 subunits of the mammalian
  soluble guanylyl cyclase genes."
RL Proc. Natl. Acad. Sci. U.S.A. 97:10878-10883(2000).
DR EMBL; AF020339; AAB94876.1; -.
DR EMBL; AF297083; AAG17447.1; -.
DR HSSP; P16068; IAWN.
DR MGI; MGI:1860604; Gucylb3.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.
KW lyase.
SQ SEQUENCE 620 AA; 70597 MW; EED89ABF539F8E9 CRC64;

```

Query Match 21.8%; Score 781.5; DB 11; Length 620;  
 Best Local Similarity 35.5%; Pred. No. 1.1e-52;  
 Matches 200; Conservative 97; Mismatches 209; Indels 57; Gaps 14;

```

OY 101 KIKESKSLERDEFKTTAEQAVAGVPEVETIKESIGEYFKICYED-ENIG 153
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 30 QLDEGQFLVRLITYDSDKTYDLYVAASRYVLNINAGEIL-QMGKMFVFCQSGSGDYTLIR 88
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 154 VVGFTLKPFLNFESESTLLKQSSHCQAGKRGRLSDSILCLDXE--DDFLHYVYFFPKRTS 212
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 VLGSNVRFLQNLAL-----HDLATITYPGRAPDSFRCTDAEKKGKGLIHLHYSEREGIQ 143
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:03:32 ; Search time 14.2322 Seconds

(without alignments)  
1426.467 Million cell updates/sec

Title: US-09-762-767a-2

Perfect score: 3593

Sequence: 1 MFCYKRLKDKITGECPFSL.....QKKDVEDGNANFLKASGID 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCUTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	407.5	11.3	1075	1	US-07-623-033-2
2	325.5	9.1	1077	4	US-09-412-210-1
3	322	9.0	1064	3	US-08-726-214-8
4	319.5	8.9	1099	3	US-08-726-214-14
5	319	8.9	1090	3	US-08-307-896-3
6	319	8.9	1090	3	US-08-726-214-4
7	319	8.9	1090	5	PCT-US95-11808-3
8	303	8.4	1144	3	US-08-726-214-6
9	302	8.4	1134	3	US-08-726-214-2
10	299.5	8.3	533	3	US-08-726-214-18
11	287.5	8.0	1253	4	US-08-864-785-2
12	284.5	7.9	1248	3	US-08-726-214-16
13	283.5	7.9	1098	3	US-08-726-214-10
14	282.5	7.9	604	4	US-09-008-097-4
15	282.5	7.9	1167	4	US-09-008-097-6
16	280.5	7.8	1261	4	US-09-473-716-2
17	279.5	7.8	1168	4	US-09-474-076-2
18	279.5	7.6	1180	3	US-08-726-214-12
19	272.5	7.6	1165	1	US-08-240-357-2
20	263.5	7.3	1294	4	US-09-473-717-2
21	260.5	7.3	1305	4	US-08-864-785-3
22	260.5	7.3	1353	3	US-08-894-173-2
23	260.5	7.3	1353	4	US-09-398-193-2
24	260.5	7.3	1353	4	US-09-473-717-3
25	255.5	7.1	1353	4	US-09-398-193-09
26	195	5.4	670	2	US-08-997-080-178
27	195	5.4	670	2	US-08-997-362-178

28	195	5.4	670	4	US-09-095-855-178	Sequence 178, App
29	195	5.4	670	4	US-09-324-542-178	Sequence 178, App
30	195	5.4	670	4	US-09-205-426-178	Sequence 178, App
31	195	5.4	722	2	US-08-997-080-174	Sequence 174, App
32	195	5.4	722	2	US-08-997-362-174	Sequence 174, App
33	195	5.4	722	2	US-09-095-855-174	Sequence 174, App
34	195	5.4	722	4	US-09-324-542-174	Sequence 174, App
35	195	5.4	722	4	US-09-205-426-174	Sequence 174, App
36	192.5	5.4	107	4	US-08-857-076-76	Sequence 76, App
37	191	5.3	112	4	US-08-857-076-75	Sequence 75, App
38	187	5.2	419	2	US-08-997-080-125	Sequence 125, App
39	187	5.2	419	2	US-08-997-362-125	Sequence 125, App
40	187	5.2	419	4	US-09-095-855-125	Sequence 125, App
41	187	5.2	419	4	US-09-324-542-125	Sequence 125, App
42	187	5.2	419	4	US-09-205-426-125	Sequence 125, App
43	180	5.0	429	3	US-08-307-896-4	Sequence 4, App
44	180	5.0	429	3	PCT-US95-11808-4	Sequence 4, App
45	138.5	3.9	85	3	US-08-894-173-59	Sequence 59, App

## ALIGNMENTS

```
RESULT 1
US-07-623-033-2
; Sequence 2, Application US/07623033
; Patent No. 5237051
GENERAL INFORMATION:
APPLICANT: Garbers, David L.
APPLICANT: Schultz, Stephanie
TITLE OF INVENTION: CLONING THE ENTEROTOXIN RECEPTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: TILTON, FALLON, LUNGMS & CHESTNUT
STREET: 100 South Wacker Drive - Suite 960
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60606-4002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/623,033
FILING DATE: 19901206
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: VU9018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 456-8000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1075 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-623-033-2
Query Match 11.3% Score 407.5; DB 1; Length 1075;
Best Local Similarity 36.6% Pred. No. 3.1e-31;
Matches 98; Conservative 47; Mismatches 94; Indels 29; Gaps 6;
QY 423 DGLKRL-----GKLTATLEQAQAALAEERKKTVDLCSIFPEVAQAOLMOGOVQAARFS 478
DB 764 DTLIRRLQLYSRNLEHLYEERTQLYKABRDADHLNFMILPRLVYKSLKEGIVPELYE 823
QY 479 NVTALFSDIVFTALSCSCSPLOYITMNLALYTRPDQCGGLDYKVTITGDAYCVAGGL 538
DB 824 EVTIYFSDIVFTICKYSTPEVVDMLNDIYKSPDQIVDHHVDYKVTETIDAVYVAGSL 883
```

RESULT 3  
US-08-726-214-8  
Sequence 8, Application US/08726214  
Patent No. 6107076  
GENERAL INFORMATION:  
APPLICANT: Tang, Wei-Jen  
APPLICANT: Gilman, Alfred G.  
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASIS  
TITLE OF INVENTION: AND USGS THEREFOR  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee

RESULT 4  
US-08-726-214-14  
; Sequence 14, Application US/08726214  
; Patent No. 6107076  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Wei-Jen  
; APPLICANT: Gilman, Alfred G.  
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE  
; TITLE OF INVENTION: AND USES THEREFOR  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: United States of America

ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/726,214  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,498  
FILING DATE: 04-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: DTSD:450  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1099 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-726-214-14

Query Match 8.9% Score 319.5; DB 3; Length 1099;  
Best Local Similarity 26.3%; Pred. No. 1.9e-22;  
Matches 82; Conservative 64; Mismatches 95; Indels 71; Gaps 8;

QY 376 SSATLFGSPCVRLDEFTGKGLYLSIDIPINHALRDVLLIGEQAQADGLKRLKAT 435  
DB 184 ANAVTLILG-----FTGA---FKHQLODASRDLEFY-----T 214  
QY 436 LE-QAQAALBEKKKVDLCSTFCEVA-----QQLMG-----QVQA 474  
DB 215 VKCLOIRKRLVERKQOENLLSVLPANISGMKLAIERKEGDRHMPDNHSLYV 274  
QY 475 KFSNVTMLFSIDYGTATSCSPLOVITMLNALTREDOQCGELDYKVEITGDYCV 534  
DB 275 KRQNVSLTADYIGFTRLASDCSPKELVYVNLNLFGRFDIAKANEOMRIKILGDCYCC 334  
QY 535 AGGLHESDTHAVQIALMALKMEISDEYMSPHGEPIKRIGLHSGSVFAGVGMKPRY 594  
DB 335 VSGLPVSLPTHARNCVMGLDICEAIKQVREATGVDISMRYGSHGNVLCGVIGLRKQY 394  
QY 595 CLEGNNTLANKEFSCVPRKINVSPTTYRL-----KDCPG-----EV 633  
DB 395 DVMSHDVSILANRMEAGVPGRVHITETALNHLDRAYEVEDHGQRDPYLEKMNIRTYLV 454  
QY 634 FTPRSRELPN 645  
DB 455 IDPRSQQPPPS 466

RESULT 5  
US-08-307-896-3  
Sequence 3, Application US/08307896C  
Patent No. 6034071  
GENERAL INFORMATION:  
APPLICANT: Iyengar, Srinivas Ravi  
TITLE OF INVENTION: MUTANT ACTIVATED GSALPHA AND ADENYLYL  
TITLE OF INVENTION: CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS  
FILE REFERENCE: 29770  
CURRENT APPLICATION NUMBER: US/08/307,896C  
CURRENT FILING DATE: 1994-09-16  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 1090  
TYPE: PRT

ORGANISM: Rattus norvegicus  
US-08-307-896-3

Query Match 8.9% Score 319; DB 3; Length 1090;  
Best Local Similarity 30.4%; Pred. No. 2.1e-22;  
Matches 80; Conservative 53; Mismatches 86; Indels 44; Gaps 7;

QY 398 LYLSDIP-----IHNALRDVLLI--GQARAQGLKRLGLKATLEQAHQ----- 441  
DB 174 VYLSATGAKKHLWQILANIYIFICGLAGA---YKHL--MELAQYTRDNCIKS 228  
QY 442 --ALEBKKTVDLCSTFCEVAQQLMGQVQ-----AKFSN 479  
DB 229 RIKLEFERKQOERILLSTLPAHIAMEM-KAEIOLQPKAQOMENTNPNFLYKRHTN 287  
QY 480 VTMLFSIDYGTATSCSPLOVITMLNALTREDOQCGELDYKVEITGDYCVAGSLH 539  
DB 288 VSLTADYIGFTRLASDCSPKELVYVNLNLFGRFDIAKANEOMRIKILGDCYCVSGLP 347  
QY 540 KESDTHAVQIALMALKMEISDEYMSPHGEPIKRIGLHSGSVFAGVGMKPRYCLFGN 599  
DB 348 ISLPNHAKNCYKMGIDMCEAIKRYRDAATGVDINNRVGHSGNVLCGVIGLRKQYDWSH 407  
QY 600 NTLANKFSCVPRKINVSPTT 622  
DB 408 DVTLANHMEAGVPGRVHISSTV 430

RESULT 6  
US-08-726-214-4  
Sequence 4, Application US/08726214  
Patent No. 6107076

GENERAL INFORMATION:  
APPLICANT: Tang, Wei-Jen  
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/726,214  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,498  
FILING DATE: 04-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: DTSD:450  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1090 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-726-214-4

Query Match 8.9% Score 319; DB 3; Length 1090;

Best Local Similarity 30.4%; Pred. No. 2,1e-22;  
Matches 80; Conservative 53; Mismatches 86; Indels 44; Gaps 7;

QY 398 LYSIDP-----IHNLROYVLI--GEORADGLKRLGKATLEQAHO-----441  
Db 174 VYLSATGAKENHLEWQILAVIIFICNLGA--YKHL--MELALQYVRCNCIKS 228  
QY 442 --ALEEKKRTVDLCSIFPCEVAOOLMOGOVVO-----AKFSN 479  
Db 229 RILFEKROERLLSLRAHAMEK-KAEIIRLOGPAGOMENTNPHNLVYKRHTN 287  
QY 480 VTMLESDIVGFTALCSQSPLOYITMLALYTRFDQCGELDYKVTIGDAYCAGGLH 539  
Db 288 VSILYADIVGFTRLASDPSGELVHMLNELFGKFDQIAKENECMRKIKLGDYCVSGLP 347  
QY 540 KESDTHAVOIALAKRMELSDVMSHPGEPIMKRIHSGSVFAGVYKMPRYCLFGN 599  
Db 348 ISLPHAKNCVKKGLDCEAIKRVKRDATVDINMRVGHSGVNLGVLGLOKQYDVWSH 407  
QY 600 NWTLANFESCSVPKINVSPTT 622  
Db 408 DVTLANHMEAGVPGRVHISSTV 430

RESULT 7  
PCT-US95-11808-3

; Sequence 3, Application PC/TUS9511808  
; GENERAL INFORMATION:  
; APPLICANT: Iyengar, Sriharas Ravi V.  
; TITLE OF INVENTION: MUTANT ACTIVATED GSALPHA AND  
; TITLE OF INVENTION: ADENITYL  
; TITLE OF INVENTION: CYCLASE 2 FOR USE AS THERAPEUTIC AGENTS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brumbaugh, Graves, Donohue and  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U S  
; ZIP: 10112-0228  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/11808  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/307,896  
; FILING DATE: 16-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Richard S.  
; REGISTRATION NUMBER: 26,154  
; REFERENCE/DOCKET NUMBER: 29970 165/28755  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 408-2500  
; TELEFAX: (212) 765-2519  
; TELEX: 650 6111063  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1090 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-11808-3

Query Match 8.9%; Score 319; DB 5; Length 1090;  
Best Local Similarity 30.4%; Pred. No. 2.1e-22;  
Matches 80; Conservative 53; Mismatches 86; Indels 44; Gaps 7;

QY 398 LYSIDP-----IHNLROYVLI--GEORADGLKRLGKATLEQAHO-----441  
Db 174 VYLSATGAKENHLEWQILAVIIFICNLGA--YKHL--MELALQYVRCNCIKS 228  
QY 442 --ALEEKKRTVDLCSIFPCEVAOOLMOGOVVO-----AKFSN 479  
Db 229 RILFEKROERLLSLRAHAMEK-KAEIIRLOGPAGOMENTNPHNLVYKRHTN 287  
QY 480 VTMLESDIVGFTALCSQSPLOYITMLALYTRFDQCGELDYKVTIGDAYCAGGLH 539  
Db 288 VSILYADIVGFTRLASDPSGELVHMLNELFGKFDQIAKENECMRKIKLGDYCVSGLP 347  
QY 540 KESDTHAVOIALAKRMELSDVMSHPGEPIMKRIHSGSVFAGVYKMPRYCLFGN 599  
Db 348 ISLPHAKNCVKKGLDCEAIKRVKRDATVDINMRVGHSGVNLGVLGLOKQYDVWSH 407  
QY 600 NWTLANFESCSVPKINVSPTT 622  
Db 408 DVTLANHMEAGVPGRVHISSTV 430

RESULT 8  
US-08-726-214-6

; Sequence 6, Application US/08726214  
; Patent No. 6107076  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Wei-Jen  
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENITYL CYCLASE  
; TITLE OF INVENTION: AND USES THEREFOR  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: United States of America  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/726,214  
; FILING DATE: Concurrently Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/005,498  
; FILING DATE: 04-OCT-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Highlander, Steven L.  
; REGISTRATION NUMBER: 37,642  
; REFERENCE/DOCKET NUMBER: UTSID:450  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (512) 474-7577  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1144 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; US-08-726-214-6

Query Match 8.4%; Score 303; DB 3; Length 1144;  
Best Local Similarity 30.7%; Pred. No. 8.8e-21;  
Matches 84; Conservative 50; Mismatches 92; Indels 48; Gaps 9;

QY 433 KATLE-----QAHQALEEKKRTVDLCSIFPCEVAOOL-----WQGVYVOA 474  
Db 253 KAFLEARQSLVEKMKNLBEGSQOQENMLSLPKHVADELMKMKDESKQKQOQVNTWYM 312





QY 614 KINVSPTTYRL 626  
 DB 203 GRVHITKTTLACT 215

## RESULT 11

US-08-864-785-2  
 ; Sequence 2, Application US/08864/785A  
 ; Patent No. 6329566  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kaplan, Joshua M.  
 ; APPLICANT: Oppenheimer, Allison J.  
 ; APPLICANT: Hart, Anne C.  
 ; TITLE OF INVENTION: METHODS FOR THE DETECTION, TREATMENT,  
 ; FILE OF INVENTION: AND PREVENTION OF NEURODEGENERATION  
 ; FILE REFERENCE: 00786/353001  
 ; CURRENT APPLICATION NUMBER: US/08/864,785A  
 ; CURRENT FILING DATE: 1997-05-29  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: FASTSEQ for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1253  
 ; TYPE: PRT  
 ; ORGANISM: Caenorhabditis elegans  
 US-08-864-785-2

Query Match 8.0%; Score 287.5; DB 4; Length 1253;  
 Best Local Similarity 30.5%; Pred. No. 3.6e-19;  
 Matches 69; Conservative 42; Mismatches 78; Indels 37; Gaps 2;

QY 439 AHQALEERKKTVDLCIFCEVAQOLMOG----- 469  
 DB 240 AKKDELETOEFQHMIGSVMPKRVADDELKDNASELRPSASNSDCRTSMATOVDOPLAK 299  
 QY 470 QVQAQKFF-----SNVTMLFSDIVGFTALCSQSPLOYITMLNATYTRDQCGELD 521  
 DB 300 MVPEKRRKRPFTMTNMTNVSILFADLAGTKMSKSKADELVNLTDLFGREDYICRLRG 359  
 QY 522 YKVTETIDGACVAGLHRESPTHAVALMALKMEISDEVMSPHGEPIKRIQLHSGS 581  
 DB 360 LEKISTLDCCYCVAGCEPCDDHACRTVEGMGLDMIVAIRQPDIDRGQEVNRRVGIHNGK 419  
 QY 582 VFAGVGVKMPRYCLFGNNVTLANKEFSGSVPRKINVSPTTYRLK 627  
 DB 420 VMCGAVGTRKRFKDFVNSDVTLANMESGVAGRVHVSATKLK 465

## RESULT 12

US-08-726-214-16  
 ; Sequence 16, Application US/08726/214  
 ; Patent No. 6107076  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Wei-Jen  
 ; APPLICANT: Gilman, Alfred G.  
 ; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE  
 ; TITLE OF INVENTION: AND USES THEREFOR  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Arnold, White & Durkee  
 ; STREET: P. O. Box 4433  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: United States of America  
 ; ZIP: 77210  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: IBM PC compatible  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/726,214  
 ; FILING DATE: Concurrently Herewith  
 ; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/005,498  
 ; FILING DATE: 04-OCT-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Highlander, Steven L.  
 ; REGISTRATION NUMBER: 37,642  
 ; REFERENCE/DOCKET NUMBER: UTSID:450  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (512) 418-3000  
 ; TELEFAX: (512) 474-7577  
 ; INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1248 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 US-08-726-214-16

Query Match 7.9%; Score 284.5; DB 3; Length 1248;  
 Best Local Similarity 22.5%; Pred. No. 7e-19;  
 Matches 147; Conservative 99; Mismatches 217; Indels 189; Gaps 31;

QY 103 KESRSLERDEPEKIAQAVAA--GVPVEYIKESIGEVEYKICEED----- 148  
 DB 568 KERNEFLKKNHNIETYLINQPEBSLSLPEDIKVESV-----CSDRRNSGATTEGSWS 621  
 QY 149 -----ENILG---VVGFTLKDFLNSFTLLKQSSHCQ-----EAGKRG 184  
 DB 622 PELPNDIVGNQNTLAALTRNSINLPHNLQALHVGSPPEINKRIEHTIDLSGDLK 681  
 QY 185 LEDASILCLXEDDFL-HVY-----YFPKRTSLIDG----- 217  
 DB 682 REHIFPSIMFSDSLERKYSOMDEVEKSNLYCAFIYLFTTAQSL-LRSSRLMPTI 740  
 QY 218 -----IIRAALVYETEVEYSIMPCFHNOCSEFVNOPYLLSYHMSSTPSPSKPQ 272  
 DB 741 QESILIMHSALVITTTEDKICPLILIRKCC-WINETYIARVNTIA----- 788  
 QY 273 SSLV-----IPTSLECKTFPHFMFKDMTI--LOFGNGIRLNNRRDFOGRPNFEYFE 325  
 DB 789 SILINFLGAVINIIMCD-----FDKSIPLKNLTFNS-----AVFTDICSYPEV 833  
 QY 326 ILTPKINTPGSGIMTMAMQFVVRVRNRDNSYKSSRYMDLKGMITYIVESSAI--LFL 382  
 DB 834 -----VFQGVLMATVCAVFLRL---NSYLKLAVALIIMA--ITYALLTETIYAGLFL 879  
 QY 383 GSPCYDRL-----EDFTG-----RGLYLSDIPILNALRDVYLIGEQARAQDGLKRLG 430  
 DB 880 S---YDNLNHSGEPLTGKREASLLIMAFILAVFH-----GQDL----- 916  
 QY 431 KIKATLE-----QAQALEERK--KKTVDLCSIFCEVAQOLMO-----GQVQAQKFS 478  
 DB 917 EYTAFLDPLMRVQAKEEINEMKDLREHNENMLRNILPHVARHFLERDORNEEYLSQYD 976  
 QY 479 NYTMLFSDIVGFTALCSQ-----SPLOYITMLNATYTRDQCGE--LDYKVTETIGDA 531  
 DB 977 AVGVAFASIPGFAFETYSQTEKNNQVECLRLINELIADFDELDEGRQDEKIKITIGST 1036  
 QY 532 YCVAGLHRESPT-----HAVQIALMALKMEISDEVMSPHGEPIKRIQLHSGSVAG 585  
 DB 1037 YMAVSGLSPEKQOCEDKRGHICALADPSLATFESIQEINKHSFNNFELRIQISHGSVAG 1096  
 QY 586 VGVGMPRYCLFGNNVTLANKEFSGSVPRKINVSPTTYRLKDCPGVFETPR 637  
 DB 1097 VIGAKKPYQDLMGKVNVLASRDSTGSGRIGVPEETYLILKD-QGFAFDPR 1147

## RESULT 13

US-08-726-214-10  
 ; Sequence 10, Application US/08726/214  
 ; Patent No. 6107076  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Wei-Jen

APPLICANT: Gilman, Alfred G.  
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE  
TITLE OF INVENTION: AND USES THEREFOR  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/726,214  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,498  
FILING DATE: 04-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: 6TSD:450  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1098 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-726-214-10

Query Match 7.9%; Score 283.5; DB 3; Length 1098;  
Best Local Similarity 26.4%; Pred. No. 7.1e-19;  
Matches 83; Conservative 52; Mismatches 108; Indels 71; Gaps 8;

QY	438	QAHQALBEEKKRYVDLCSTFPCVAAQOL-----WQGVYQAKKFSNVTMLFSDI	487
DB	253	QARHSQRENQOERLLSVLPBRVAMEMKADINAKORDMFFHIYIOKHNVSTILFADI	312
QY	488	VGFPAISQCSPLQVITMNLALYTRPDQCGELDYVKVETIGDAYCVAGSLHKESDTHAV	547
DB	313	EGFTSLASQCTAOELVMTLNELFARFDKLAENHCLRIKILGDCYCVSGLEPARADAH	372
QY	548	QIALMALKMMELSDVMSPHGEPIKRIKIGHSVYFAGVGVKMPRYCLFGNNVTLANKE	607
DB	373	CCVEMGMDIMEAISVREVTGVNVMRGVSHSGRVGCVGLGRKMQFVWSNDVTLANHM	432
QY	608	ESGCVPRKINVSPT-----YRLKDCPG-----FVF---TPRSSEEL--	642
DB	433	EAGKAGRIHITKATLYNLNGDYVEVPCGGERNAYLKEHSIEYFLIRCTOKREERAM	492
QY	643	-----PNPFSIEPGICHFILAYQOGTSKPCFOKK-----YVEDGNA---	680
DB	493	IARKNRORTNSIGNPFRHGAERFFYNHL-----GGNQVSKEMKMGFEDPKDNOAES	546
QY	681	-----NFLGKA	686
DB	547	ANPEDEVDELGRA	560

RESULT 14  
US-09-008-097-4  
Sequence 4, Application US/09008097  
Patent No. 6306830  
GENERAL INFORMATION:  
APPLICANT: Hammond, H. Kirk

APPLICANT: Insel, Paul A.  
APPLICANT: Ping, Peipel  
APPLICANT: Post, Steven R.  
APPLICANT: Gao, Melina  
TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE  
HEART FAILURE  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/008,097  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dylan, Tyler M.  
REGISTRATION NUMBER: 37,612  
REFERENCE/DOCKET NUMBER: 22000-20567.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-813-5600  
TELEFAX: 650-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 604 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-09-008-097-4

Query Match 7.9%; Score 282.5; DB 4; Length 604;  
Best Local Similarity 31.8%; Pred. No. 3.3e-19;  
Matches 63; Conservative 44; Mismatches 82; Indels 9; Gaps 1;

QY	438	QAHQALBEEKKRYVDLCSTFPCVAAQOLMOG-----QVYQAKKFSNVTMLFSDIV	488
DB	115	QARLHLQHNHQOERLLSVLPQHVAMEMKEDINTKEDMFFHIYIOKHNVSTILFADIE	174
QY	489	GFTALCSQCSPLQVITMNLALYTRPDQCGELDYVKVETIGDAYCVAGSLHKESDTHAV	548
DB	175	GFTSLASQCTAOELVMTLNELFARFDKLAENHCLRIKILGDCYCVSGLEPARADAH	234
QY	549	IATMALKMMELSDVMSPHGEPIKRIKIGHSVYFAGVGVKMPRYCLFGNNVTLANKE	608
DB	235	CEVEMGVDMIMEAISVREVTGVNVMRGVSHSGRVGCVGLGRKMQFVWSNDVTLANME	294
QY	609	SCSVPRKINVSPTTYRL	626
DB	295	AGSRAGRITRATLOYL	312

RESULT 15  
US-09-008-097-6  
Sequence 6, Application US/09008097  
Patent No. 6306830  
GENERAL INFORMATION:  
APPLICANT: Hammond, H. Kirk  
APPLICANT: Insel, Paul A.

APPLICANT: Post, Steven R.  
APPLICANT: Gao, Melina  
TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE  
TITLE OF INVENTION: HEART FAILURE  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & ROEBSTER  
STREET: 755 PAGE MILL ROAD  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/008,097  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dylan, Tyler M  
REGISTRATION NUMBER: 37,612  
REFERENCE/DOCKET NUMBER: 22000-20567, 21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-813-5600  
TELEFAX: 650-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1167 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-09-008-097-6

Query Match 7.9% Score 282.5; DB 4; Length 1167;  
Best Local Similarity 31.8%; Pred. NO. 9.9e-19;  
Matches 63; Conservative 44; Mismatches 82; Indels 9; Gaps 1;  
QY 438 QAHQALBEKKKKTVDLCSIPCEVAOOLMOG-----QVQAKKFSNVTMLFSDIV 488  
DB 326 QARHLQEHENQGERLLSVLPQHVAMEMKEDINTKKEDMFKIYIQKHDNVSTILFADIE 385  
QY 489 GFTAICSOCSPLQVITMIALYTRFDQCGELDVYKVTIGDAYCVAGGLHKESDTHAVQ 548  
DB 386 GFTSLASOCTAOELVMTLNEFARFDKLAENHCLRIKILGDCYCVSGLPEARADRAHC 445  
QY 549 IALNALKMMELSDVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNNTLANKE 608  
DB 446 CVERKGVDTIEAISTLVREYTGVMNMRVGIHSGRVHCGVILGRKWFVMSNDVTLANME 505  
QY 609 SCSVPRKINVSPTTYRL 626  
DB 506 AGSRAGRHIHTRATLTQYL 523

Search completed: June 27, 2003, 13:05:57  
Job time: 16.2322 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: June 27, 2003, 13:05:32 ; Search time 36.3713 Seconds  
(Without alignments)  
2080.112 Million cell updates/sec

Title: US-09-762-767a-2  
Perfect score: 3593  
Sequence: 1 MCTWKIKDKITGECPFSL.....QKKQVEDGNANFLGKASGID 690

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	3593	100.0	9 US-10-205-823-162	Sequence 162, App
2	3265.5	90.9	9 US-09-952-2130-2	Sequence 2, Appl
3	408.5	11.4	10 US-10-157-031-18	Sequence 18, Appl
4	408.5	11.4	10 US-09-819-249-2	Sequence 2, Appl
5	325.5	9.1	10 US-10-121-911-1	Sequence 1, Appl
6	316	8.8	10 US-10-282-942-2	Sequence 2, Appl
7	303.5	8.4	10 US-09-925-297-811	Sequence 811, App
8	282.5	7.9	10 US-09-750-240-4	Sequence 4, Appl
9	282.5	7.9	10 US-09-750-240-6	Sequence 6, Appl
10	280.5	7.8	9 US-10-175-158-2	Sequence 2, Appl
11	279.5	7.8	9 US-10-201-000-2	Sequence 2, Appl
12	279.5	7.8	10 US-09-750-240-11	Sequence 11, Appl
13	278	7.7	10 US-09-750-240-13	Sequence 13, Appl
14	263.5	7.3	12 US-10-071-223-2	Sequence 2, Appl
15	260.5	7.3	10 US-09-751-1008-2	Sequence 2, Appl
16	260.5	7.3	12 US-10-071-223-3	Sequence 9, Appl
17	255.5	7.1	1353 9 US-09-751-1008-99	Sequence 95, Appl
18	212	5.9	257 9 US-09-989-442-95	Sequence 794, App
19	207	5.8	241 9 US-09-764-868-794	

20	207	5.8	241 9 US-09-989-442-141	Sequence 141, App
21	195	5.4	670 9 US-10-051-643-178	Sequence 178, App
22	195	5.4	670 9 US-09-880-505-178	Sequence 178, App
23	195	5.4	722 9 US-10-051-643-174	Sequence 174, App
24	195	5.4	722 9 US-09-880-505-174	Sequence 174, App
25	194	5.4	302 10 US-09-915-582-50	Sequence 50, Appl
26	192.5	5.4	107 10 US-09-205-658-76	Sequence 76, Appl
27	192.5	5.4	107 10 US-09-844-353A-76	Sequence 76, Appl
28	191	5.3	112 10 US-09-205-658-75	Sequence 75, Appl
29	191	5.3	112 10 US-09-844-353A-75	Sequence 75, Appl
30	187	5.2	419 9 US-10-051-643-125	Sequence 125, App
31	187	5.2	419 9 US-09-880-505-125	Sequence 125, App
32	181.5	5.1	311 10 US-09-925-297-515	Sequence 515, App
33	174	4.8	225 9 US-09-989-442-92	Sequence 92, Appl
34	148	4.1	109 9 US-09-764-868-795	Sequence 92, Appl
35	148	4.1	109 9 US-09-989-442-143	Sequence 143, App
36	148	4.1	109 9 US-10-091-504-1060	Sequence 1060, App
37	148	4.1	109 10 US-09-764-868-1060	Sequence 1060, App
38	138.5	3.9	85 10 US-09-751-1008-59	Sequence 59, Appl
39	137.5	3.8	85 10 US-09-751-1008-60	Sequence 60, Appl
40	137.5	3.8	85 10 US-09-751-1008-61	Sequence 61, Appl
41	133.5	3.7	85 10 US-09-751-1008-54	Sequence 54, Appl
42	133.5	3.7	85 10 US-09-751-1008-55	Sequence 55, Appl
43	133.5	3.7	85 10 US-09-751-1008-56	Sequence 56, Appl
44	133.5	3.7	85 10 US-09-751-1008-57	Sequence 57, Appl
45	133.5	3.7	85 10 US-09-751-1008-58	Sequence 58, Appl

## ALIGNMENTS

RESULT 1  
US-10-205-823-162  
; Sequence 162, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarepu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangt  
; APPLICANT: Monsey, Angela M.  
; APPLICANT: Glaty, Karen  
; APPLICANT: Zhao, Xuneli  
; APPLICANT: Anderson, Dustin  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; FILE REFERENCE: MRI-044  
; CURRENT APPLICATION NUMBER: US/10/205, 823  
; CURRENT FILING DATE: 2002-07-25  
; PRIOR APPLICATION NUMBER: 60/307, 982  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: 60/314, 356  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/325, 020  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: 60/341, 746  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/362, 158  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 455  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 162  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-205-823-162  
Query Match 100.0%; Score 3593; DB 9; Length 690;  
Best Local Similarity 100.0%; Pred. No. 1e-305;

Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCTRLKDKLITGECFSLAPGVNNESSSEAAAGSSSSCKATVPICODIPKNIQESLP 60  
 Db 1 MCTRLKDKLITGECFSLAPGVNNESSSEAAAGSSSSCKATVPICODIPKNIQESLP 60  
 QY 61 ORKTSRSRVYLHTLAESICKLIFPEFERLNALORTLAKHKIKESRKSLEDEDEKTIAE 120  
 Db 61 ORKTSRSRVYLHTLAESICKLIFPEFERLNALORTLAKHKIKESRKSLEDEDEKTIAE 120  
 QY 121 QAVAGVVEVYIKESLGEVEFKICEEDENILGVVGGTLKDFLNSFTLLKQSSHCQAG 180  
 Db 121 QAVAGVVEVYIKESLGEVEFKICEEDENILGVVGGTLKDFLNSFTLLKQSSHCQAG 180  
 QY 181 KRGLEDASILCLDKEDDFLHVYFFPKRTSLIPGIIKAAHVLTEVEVSLMPCE 240  
 Db 181 KRGLEDASILCLDKEDDFLHVYFFPKRTSLIPGIIKAAHVLTEVEVSLMPCE 240  
 QY 241 HNDSEFVNOPYLLYSVHMKSTKPSLSPKPOSSLVPTSLFCKTFPHFMDKMTILQ 300  
 Db 241 HNDSEFVNOPYLLYSVHMKSTKPSLSPKPOSSLVPTSLFCKTFPHFMDKMTILQ 300  
 QY 301 FNGGIRRLMNRDPOGKPNFEFEIILPKINOTFSGIMTMNQFVYVRVRMDNSYKKS 360  
 Db 301 FNGGIRRLMNRDPOGKPNFEFEIILPKINOTFSGIMTMNQFVYVRVRMDNSYKKS 360  
 QY 361 SRVMDLKQMIYIYESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNAIRDVYLIGEQAR 420  
 Db 361 SRVMDLKQMIYIYESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNAIRDVYLIGEQAR 420  
 QY 421 AODGLKRLGLKATLEAHQALBEKKRTVDLCSIFPCVAAQOLMOGOVAAKKEFNV 480  
 Db 421 AODGLKRLGLKATLEAHQALBEKKRTVDLCSIFPCVAAQOLMOGOVAAKKEFNV 480  
 QY 481 TMLFSDIYGFALISQCSPLQVITMLNLYTRFDQCGELDYVETIGDAYCYVAGGLHK 540  
 Db 481 TMLFSDIYGFALISQCSPLQVITMLNLYTRFDQCGELDYVETIGDAYCYVAGGLHK 540  
 QY 541 ESDTHAVOIALMALKMELSDVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 Db 541 ESDTHAVOIALMALKMELSDVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 QY 601 VTLANKFSCSVPRKINVSPTTYRLKDCPGFVTPRSRELPNFPSEIIGICHFLDAY 660  
 Db 601 VTLANKFSCSVPRKINVSPTTYRLKDCPGFVTPRSRELPNFPSEIIGICHFLDAY 660  
 QY 661 QOGTNSKPCFOKKDVEDGNANFLGKASGID 690  
 Db 661 QOGTNSKPCFOKKDVEDGNANFLGKASGID 690

RESULT 2  
 US-09-952-213D-2  
 ; Sequence 2, Application US/09952213D  
 ; Publication No. US20030096240A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MURAD, FERID  
 ; APPLICANT: SHARINA, IRAIDA G.  
 ; APPLICANT: KROMENACKER, J. S.  
 ; APPLICANT: MARTIN, E.  
 ; TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC  
 ; FILE REFERENCE: UTSH:252US  
 ; CURRENT APPLICATION NUMBER: US/09/952,213D  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 691  
 ; TYPE: PRF  
 ; ORGANISM: Mus musculus  
 US-09-952-213D-2

Query Match 90.9%; Score 3265.5; DB 9; Length 691;

Best Local Similarity 89.6%; Pred. No. 4.8e-277;  
 Matches 619; Conservative 37; Mismatches 34; Indels 1; Gaps 1;

QY 1 MCTRLKDKLITGECFSLAPGVNNESSSEAAAGSSSSCKATVPICODIPKNIQESLP 60  
 Db 1 MCTRLKDKLITGECFSLAPGVNNESSSEAAAGSSSSCKATVPICODIPKNIQESLP 60  
 QY 61 ORKTSRSRVYLHTLAESICKLIFPEFERLNALORTLAKHKIKESRKSLEDEDEKTIAE 120  
 Db 61 ORKTSRSRVYLHTLAESICKLIFPEFERLNALORTLAKHKIKESRKSLEDEDEKTIAE 120  
 QY 121 QAVAGVVEVYIKESLGEVEFKICEEDENILGVVGGTLKDFLNSFTLLKQSSHCQAG 180  
 Db 121 QAVAGVVEVYIKESLGEVEFKICEEDENILGVVGGTLKDFLNSFTLLKQSSHCQAG 180  
 QY 181 KRGLEDASILCLDKEDDFLHVYFFPKRTSLIPGIIKAAHVLTEVEVSLMPCE 240  
 Db 181 KRGLEDASILCLDKEDDFLHVYFFPKRTSLIPGIIKAAHVLTEVEVSLMPCE 240  
 QY 241 HNDSEFVNOPYLLYSVHMKSTKPSLSPKPOSSLVPTSLFCKTFPHFMDKMTILQ 300  
 Db 241 HNDSEFVNOPYLLYSVHMKSTKPSLSPKPOSSLVPTSLFCKTFPHFMDKMTILQ 300  
 QY 301 FNGGIRRLMNRDPOGKPNFEFEIILPKINOTFSGIMTMNQFVYVRVRMDNSYKKS 360  
 Db 301 FNGGIRRLMNRDPOGKPNFEFEIILPKINOTFSGIMTMNQFVYVRVRMDNSYKKS 360  
 QY 361 SRVMDLKQMIYIYESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNAIRDVYLIGEQAR 420  
 Db 361 SRVMDLKQMIYIYESSAILFLGSPCVDRLEDFTGRGLYLSDIPIHNAIRDVYLIGEQAR 420  
 QY 421 AODGLKRLGLKATLEAHQALBEKKRTVDLCSIFPCVAAQOLMOGOVAAKKEFNV 480  
 Db 421 AODGLKRLGLKATLEAHQALBEKKRTVDLCSIFPCVAAQOLMOGOVAAKKEFNV 480  
 QY 481 TMLFSDIYGFALISQCSPLQVITMLNLYTRFDQCGELDYVETIGDAYCYVAGGLHK 540  
 Db 481 TMLFSDIYGFALISQCSPLQVITMLNLYTRFDQCGELDYVETIGDAYCYVAGGLHK 540  
 QY 541 ESDTHAVOIALMALKMELSDVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 Db 541 ESDTHAVOIALMALKMELSDVMSPHGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNN 600  
 QY 601 VTLANKFSCSVPRKINVSPTTYRLKDCPGFVTPRSRELPNFPSEIIGICHFLDAY 660  
 Db 601 VTLANKFSCSVPRKINVSPTTYRLKDCPGFVTPRSRELPNFPSEIIGICHFLDAY 660  
 QY 661 -QOGTNSKPCFOKKDVEDGNANFLGKASGID 690  
 Db 661 HQGPNSKPFWDKDEDGNANFLGKASGVD 691

RESULT 3  
 US-10-157-031-18  
 ; Sequence 18, Application US/10157031  
 ; Publication No. US20030108890A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baranova, A. V.  
 ; APPLICANT: Yanovsky, N. K.  
 ; APPLICANT: Kozlov, A. P.  
 ; APPLICANT: Lobashev, A. V.  
 ; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequence  
 ; FILE REFERENCE: 2760-103  
 ; CURRENT APPLICATION NUMBER: US/10/157,031  
 ; NUMBER OF SEQ ID NOS: 415  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 18  
 ; LENGTH: 1073  
 ; TYPE: PRF  
 ; ORGANISM: Homo sapiens  
 US-10-157-031-18

```

Query Match      11.4%: Score 408.5; DB 9; length 1073;
Best Local Similarity 37.3%: Pred. No. 1.2e-26;
Matches 101; Conservative 43; Mismatches 92; Indels 35; Gaps 6

QY      423 DGLKRL-----GKLTATLEAOAHOLEEKKRTVDLCSIFPCENVAQOLMOGVVOAKRFS 478
        ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db       762 DTLIRBRDQLYSRNHEHLVEERTQLYKABRDADRNLNMLPRPLVYKSLEKEGFPELXE 821
        ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY      479 NVWMLFSDIVGFTALICSQCSPLOYITMNLAYTRFDQOCGEIDVYKVFETIGDAYCVAGGL 538
        ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db       822 EVTIYFSDIVGFTICKSTPMEEVDMNDIYKSDHIVDDHDYKVETIGDAYVVASGL 881
        ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY      539 HKES-DTNAVQIALMALKMELSDVMSPH--GEIKRKRIGHSSSVAGVGYGMAPRC 595
        ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db       882 PKRNGNRHAIDIAMKALEILSFMTGELEHLPGDIWTIRIGHSPCAAGVGIMPPRYC 941
        ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY      596 LFGNNVTLANKEESCVPKRINVSPTVTRLKDCGFVFTPRSRELPNPFSEISRGCH 655
        ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db       942 LFQGITVNTASRMESTGLEPRLRHVSGSTAILLK-----RTE-----CQ 978
        ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY      656 FL-----DAYOQTNSKPCFOKRVDEEDGNAN 681
        ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db       979 FLYEVREGTYLKGKRGNETTYIWLTKGKDOKFN 1009

```

```

RESULT 4
US-09-819-249-2
Sequence 2, Application US/09819249
Patent No. US20010029019A1
GENERAL INFORMATION:
APPLICANT: Waldman, Scott A.
APPLICANT: Park, Jason
APPLICANT: Schulz, Stephanie
TITLE OF INVENTION: Compositions And Methods For Identifying And Targeting Cancer Cells
FILE REFERENCE: TUD2412
CURRENT APPLICATION NUMBER: US/09/819, 249
CURRENT FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: 60/1792, 229
PRIOR FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1073
TYPE: prt
ORGANISM: Homo sapiens
US-09-819-249-2

```

Query Match	11.4%	Score 408.5	DB 10	Length 1073
Best Local Similarity	37.3%	Pred No. 1.2e-26		
Matches 101	Conservative 43	Mismatches 92	Indels 35	Gaps 6
QY	423	DGLKKRL-----GKLKATLEAOAHQALEEKKKIVDLCSIFPEVAQOLQOGVQVQAKKFS	478	
Db	762	DTLIRRLDLYSNLNLEHYEERTQYLKARDRDRNLFTMLPRLYVKSLSKEKGFVEBELXE	821	
QY	479	NTVMLESDIVGFETALCSQCSPLQVITMNLATYRFDQOGELDYYKVFETIGDAYCVAGSL	538	
Db	822	EVTIVFSDIVGFETTCIKYSTPEVVDMLNDIYKSPDHYIDHHDDYKVFETIGDAYVYASGL	881	
QY	539	HKES-DTIAVQVIALMALKMELSDEVMSPH--GEPIKARIGLHSGSVFAGVGVKMPRYC	595	
Db	882	PRKRNRRHAIDIAKMALEILISPMGFEEHLHDPGLPIWIRIGVHSGCAGVGVGIIKMPRC	941	
QY	536	LEGNNVITLANKFESSVPRKIVNSPTTYTRRLKDCGFFVTPPSRELPNPFSEIRPICH	655	
Db	942	LEGDTVNTASRRESIGLPLRIHVSGSTIAIILK-----RTE-----CQ	978	
QY	656	FL-----DAYOOGTNSKPOCFKKDDVEDGNAN	681	
Db	979	FLIYEVKGETYTLKGRGNETTITWLTGKKDQKFN	1009	

```

RESULT 5
US-10-121-911-1
: Sequence 1, Application US/10121911
: Patent No. US20020164632A1
: GENERAL INFORMATION:
: APPLICANT: Kapeller-Libermann, Rosana
: TITLE OF INVENTION: 21529, A NOVEL ADENYLYATE CYCLASE
: FILE REFERENCE: 5800-47
: CURRENT APPLICATION NUMBER: US/10/121, 911
: CURRENT FILING DATE: 2002-04-12
: PRIOR APPLICATION NUMBER: US/09/412, 210
: PRIOR FILING DATE: 1999-10-05
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1077
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-121-911-1

```

	Query Match	9.1%;	Score 325.5;	DB 9;	length 1077;
	Best Local Similarity	28.2%;	Pred. No. 2.2e-19;		
	Matches	91;	Conservative	55;	Mismatches 98; Indels 79; Gaps 9
QY	365	DLKGMIYIVESSAILFLGSPCVDRLEDTFGGLYLSIDPIHNALRDVYLIGEOBARADG	424		
		: : : :    :    :			
Db	165	DSRPALLPOLANNAVFLCGNVA-----GVY-----HKALKE-----RA-----	198		
QY	425	LKRLGKLAKTEQA-----HQALEEEKKTVDLDCSTFP-----CEVAQOLMOGQ	470		
		: : : :    :    :			
Db	199	-----LRATFRRLSLSHSRRRRLDTEKHOBHLLSTLPAYLAREKMAEIRARLQAQ	251		
QY	471	-----VQAKKFSNYTMLFSDIYGFPAICGSPLOYITMLNLTYRFDDOCGE	519		
		: : : :    :    :			
Db	252	GSRPSTNNFHSLLYXKRHGVSVLADYIGFTRLSECSPKELVTLNLFEGFDDIAKE	311		
QY	520	LDVYEVETIGDAYCVAGGLHKESDTHAVOIALMALKMLLSDEVMSPHGEPIKMTGLHS	579		
		: : : :    :    :			
Db	312	HECMKIKILGDCGYCVSGPLPLDPHALINCVMGMGLDMCAIKRLRAATGVDIMMRGVHS	371		
QY	580	GSVFRGVYGVKMPRYCLFGNNNTLANKEFSCSVPRKINVSPTTYRLKDC-----	639		
		: : : :    :    :			
Db	372	GSVLGCVIGLQFMQDVMWSDHYTLNHHMEAGVPGRVHITGTATALLAAAYAVEDAGMEH	431		
QY	630	-----PGF-VETPRSRRE	641		
		: : : :    :    :			
Db	432	RDPYLRLEGEPTLYVIDPRAEE	454		

```

RESULT 6
US-10-282-942-2
; Sequence 2, Application US/10/282942
; Publication NO. US20030087295A1
; GENERAL INFORMATION:
; APPLICANT: Silos-Santiago, Imaculada
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF PAIN DISORDERS USING 9805E
; FILE REFERENCE: MP/01-240P1RM
; CURRENT APPLICATION NUMBER: US/10/282, 942
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 60/335, 047
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: fastseq, for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 1080
; TYPE: prt
; ORGANISM: Homo sapien
US-10-282-942-2

Query Match      8.8%; Score 316; DB 9; Length 1080;
Best Local Similarity 28.2%; Pred. No. 1,5e-18;

```

	Matches	70;	Conservative	53;	Mismatches	85;	Indels	40;	Gaps	3.
QY	438	QAHQALEEKRRKTYDLCISFFPCFVAQOOLMGQV-----	VAQKFFS	478						
Db	217	QIRRLRLRIEKKOEWELTSLVPAHLSMKMLALIERLKEHGDRCMPPNNHSLVYKHHQ	---	276						
QY	479	NYTMLSDIVGFTALCSQSCPQVITYTMALATYTRFDQCGELDYVKVETIGDAYCVAGGL	---	538						
Db	277	NSIILYADIVGFTOLASDCSPRELVVYVYVNEIFGKFDQIAKANECKRIKILGDGYCVYCSGL	---	336						
QY	539	HKESDTHVQVQALNALKMMELSDDEWMSHGEEIKRIRILSHSGSVYAGVYVKMPRYCLFG	---	598						
Db	337	FVSLPTIARNCVYKGLDMCOAIKQVREATGVDINNRVGHSGNVLCGVIGLKKMWDWMS	---	396						
QY	599	NNVTLANKFESCSVPKRLINVSPTTYRLT-----KDCPG-----	FVETPR	637						
Db	397	HDVSLANMEAGAVGVRVHIETATLKHLDKAYEVEDGHGQQRDPYLKEMNIRTYLVIDPR	---	456						
QY	638	SREELPPN	645							
Db	457	SOQPEPPS	464							

```

US-09-925-297-811
RESULT 7
: Sequence 811, Application US/09925297
: Patent No. US2002081659A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic acids, Proteins and Antibodies
: FILE REFERENCE: PA105
: CURRENT APPLICATION NUMBER: US/09/925,297
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05989
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 928
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 811
: LENGTH: 855
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (479)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (829)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-811

```

Query Match	8.48;	Score 303.5;	DB 10;	Length 855;
Best Local Similarity	29.98;	Pred. No. 1.3e-17;		
Matches 82;	Conservative 52;	Mismatches 93;	Indels 47;	Gaps 8

QY	433	TALE-----	QAHQALEEKKTAVLLCSIFCEVAQOL-----	WQGVQVA	474
Db	7	KAFLEARSL	EYKMYLEEDSOQOENIMLSILFKHVADEMLKMKDKDESQDOQOFTTMM	66	
QY	475	KKESNVTML	ESDVGFTALICSCSPLOYITMLNATLYREDOOGGELDYVKVETIGDAYCV	534	
Db	67	YRHEVSYIL	FADIVGFTOLSSACSAQELVILLNELFARBDKLAAYHQIRIKLIDGCYYC	126	
QY	535	AGGLHKEDT	HAVALMALKMMELSDENMNSHGERIKRIGLHSSVYAGVYGVAMPRY	594	
Db	127	ICGLPDYRED	HAVCISLIMGLAWEAIVYREKTKTGVDRAVGVHGTAVGLVQGORWCY	186	
QY	595	CLFGNNTLV	ANKFESCSVPRKINVAPTTYRLKDCPGFVTPRSRELPNPFSEIIG--	652	
Db	187	DYMSDVTY	AVANKMEGGLPRVHIGOST-----MDC-----	LAGEDVE--PGDG	229
QY	653	--ICHFLDAY	QOGTN-----SKPCFQKRDVEDG	678	

```

      | :| : :| ||| :| : :|
Db 230 GSRCDYLE--EKGIETVLIASKPEVKKTATQNG 261

```

```

RESULT 8
US-09-750-240-4
Sequence 4, Application US/09750240
Patient No. US20020103147A1
GENERAL INFORMATION:
APPLICANT: Hammon, H. K.
APPLICANT: Insel, P. A.
APPLICANT: Ping, P.
APPLICANT: Post, S. R.
APPLICANT: Gao, M.
TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART
TITLE OF INVENTION: FAILURE
FILE REFERENCE: 220002056723
CURRENT APPLICATION NUMBER: US/09/750,240
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 09/472,667
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: US 09/008,097
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: US 08/924,757
PRIOR FILING DATE: 1997-09-05
PRIOR APPLICATION NUMBER: US 60/048,933
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: US 08/708,661
PRIOR FILING DATE: 1996-09-05
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Fastseq for Windows Version 4.0
SEQ ID NO. 4
LENGTH: 604
TYPE: prt
ORGANISM: Homo sapiens
US-09-750-240-4

```

Query Match	7.9%;	Score 282.5;	DB 10;	Length 604;
Best Local Similarity	31.8%;	Pred. NO. 5.3e-16;		
Matches 63;	Conservative 44;	Mismatches 82;	Indels 9;	Gaps 1;

```

QY 436 QAHQALBEEKKKTVDLLCSIFPCBEVAQDLMQG-----QVYQAKFSNVLMLESIDV 488
Db 115 QALHMLQHNNRQOERILLTVLPQHYAMKMKEDINTKKEDMFKIYIKRHNVSLIFADIE 174
QY 489 GFPAISQOSPLOVITMLNATYTRFDQOCGELDYYKVETIGDAYCVAGSLHKESDFAVQ 548
Db 175 GFSLASQCTAOBLVMTLNELPARDKLAERHCLRIITLDCYCYCSGLPEARADHAC 234
QY 549 IALMALKMKELSDVYMSPHGEPIKMRIGLHSGSVFAGVYGVKMBRYCLFGNNVTLANKFE 608
Db 235 CVMGQDMLEALSLREYTVGVVNNRIVGISHGRVHCGVLGRKQDFWMSNDVTLANHME 294
QY 609 SCSVPRKINSPTYTRL 626
Db 295 AGSRAGRHIITRATLOYL 312

```

RESULT 9  
US-09-750-240-6  
Sequence 6, Application US/09750240  
Patent No. US20020103147A1  
GENERAL INFORMATION:  
APPLICANT: Hammon, H. K.  
APPLICANT: Insel, P. A.  
APPLICANT: Ping, P. A.  
APPLICANT: Post, S. R.  
APPLICANT: Gao, M.  
TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART  
FAILURE  
FILE REFERENCE: 220002056723  
CURRENT APPLICATION NUMBER: US/09/750,240  
CURRENT FILING DATE: 2001-10-12



PRIOR APPLICATION NUMBER: US 09/472,667  
PRIOR FILING DATE: 1999-12-27  
PRIOR APPLICATION NUMBER: US 09/008,097  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: US 08/924,757  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: US 60/048,933  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: US 08/708,661  
PRIOR FILING DATE: 1996-09-05  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 1167  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-750-240-6

Query Match  
Best Local Similarity 31.8%; Score 282.5; DB 10; Length 1167;  
Matches 63; Conservative 44; Mismatches 82; Indels 9; Gaps 1;

QY 438 QAHQALEEKKKTVDDLCSIFPCEVAQOLMOG-----QVYQAKKFSNTMTLFSDIY 488  
DB 326 QARLHQHNRKQERLLSVLPQHVAMEKEDINTKEDMFKHTYIQKHNDVSLFPADIE 385  
QY 489 GFTFATCSQSPLOVITMLNATYTRFDQCGELDYKVEITGDAYCVAGGLHKSPTTHAV 548  
DB 386 GFTSLASQCTAQAEIWMELNLFARFDKLAENHCLRIKILGDCYCVSGLEPARADHAHC 445  
QY 549 IALMALKMEISDEWSPHGEPIKRIKIGHSVFGAGVGVKMPRCLEGNNTLANKE 608  
DB 446 CCEMGMDMEIALSVREYGVNVMRGVHSGHGVGLGRKMOFDVMSNDVTLANHME 505  
QY 609 SCSVPKINVSPTTYRL 626  
DB 506 AGSRAGRHHITRATLOYL 523

RESULT 10  
US-10-175-158-2  
Sequence 2, Application US/10175158  
Publication No. US20030008371A1  
GENERAL INFORMATION:  
APPLICANT: Tomlinson, James  
TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL  
CYCLASE  
FILE REFERENCE: 44481-5027-01-US  
CURRENT APPLICATION NUMBER: US/10/175,158  
PRIOR FILING DATE: 2002-06-20  
PRIOR APPLICATION NUMBER: US/09/473,716  
PRIOR FILING DATE: 1999-12-29  
PRIOR APPLICATION NUMBER: PCT/US98/13540  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/070,901  
PRIOR FILING DATE: 1997-07-01  
PRIOR APPLICATION NUMBER: 08/886,362  
PRIOR FILING DATE: 1997-07-01  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1261  
TYPE: PRT  
ORGANISM: human type V adenylyl cyclase  
US-10-175-158-2

Query Match  
Best Local Similarity 7.8%; Score 280.5; DB 9; Length 1261;  
Matches 83; Conservative 52; Mismatches 108; Indels 71; Gaps 8;

QY 438 QAHQALEEKKKTVDDLCSIFPCEVAQOL-----WQGVQAKKFSNTMTLFSDI 487

DB 416 QARLHORENQOERLLSVLPQHVAMEKADINAOEDMFKHTYIQKHNDVSLFPADI 475  
QY 488 GFTFATCSQSPLOVITMLNATYTRFDQCGELDYKVEITGDAYCVAGGLHKSPTTHAV 547  
DB 476 GFTSLASQCTAQAEIWMELNLFARFDKLAENHCLRIKILGDCYCVSGLEPARADHAHC 535  
QY 548 IALMALKMEISDEWSPHGEPIKRIKIGHSVFGAGVGVKMPRCLEGNNTLANKE 607  
DB 536 CCEMGMDMEIALSVREYGVNVMRGVHSGHGVGLGRKMOFDVMSNDVTLANHME 595  
QY 608 SCSVPKINVSPTTYRL 626  
DB 596 EAGGKAGRIHITATLNTLYNGDYEBPCGGERNATYKESITFLTRCTQRKKEKAM 655  
QY 643 -----PNEPSEIPIGCHFLDAYQOGTNRKPCQK-----DVEDGNA--- 680  
DB 656 IAKMNRQRTNSIGHNPPHCAENRPFYNHL-----GQNVSKMKRKGFPDPDKNAQES 709  
QY 681 -----NFLGRA 686  
DB 710 ANPEDEYDEFLGRA 723

RESULT 11  
US-10-201-000-2  
Sequence 2, Application US/10201000  
Publication No. US20020187540A1  
GENERAL INFORMATION:  
APPLICANT: Tomlinson, James E.  
TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL  
CYCLASE  
FILE REFERENCE: 44481-5028-01-US  
CURRENT APPLICATION NUMBER: US/10/201,000  
PRIOR FILING DATE: 2002-07-24  
PRIOR APPLICATION NUMBER: US/09/474,076  
PRIOR FILING DATE: 1999-12-12  
PRIOR APPLICATION NUMBER: PCT/US98/13694  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/070,904  
PRIOR FILING DATE: 1997-07-01  
PRIOR APPLICATION NUMBER: 08/886,550  
PRIOR FILING DATE: 1997-07-01  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1168  
TYPE: PRT  
ORGANISM: human type VI adenylyl cyclase  
US-10-201-000-2

Query Match  
Best Local Similarity 7.8%; Score 279.5; DB 9; Length 1168;  
Matches 63; Conservative 47; Mismatches 82; Indels 15; Gaps 2;

QY 430 GKLTATLEQAHQALEEKKKTVDDLCSIFPCEVAQOLMOG-----QVYQAKKFSN 479  
DB 323 GYIQRHLHQHNRKQER-----LLSVLPQHVAMEKEDINTKEDMFKHTYIQKHNDV 377  
QY 480 VTMFLSDIVGFTFATCSQSPLOVITMLNATYTRFDQCGELDYKVEITGDAYCVAGGLH 539  
DB 378 VSLIFADIEGFTSLASQCTAQAEIWMELNLFARFDKLAENHCLRIKILGDCYCVSGLP 437  
QY 540 KESDTHAVQIALMALKMEISDEWSPHGEPIKRIKIGHSVFGAGVGVKMPRCLEGNNTLANKE 599  
DB 438 EAPRADHHCCEVGVNDIEIALSVREYGVNVMRGVHSGHGVGLGRKMOFDVMSNDVTLANH 497  
QY 600 NVTLANKEFSCSVPRKINVSPTTYRL 626  
DB 498 DVTLANHMEAGRAGRHHITRATLOYL 524

RESULT 12







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 27, 2003, 12:42:07 ; Search time 32.6287 Seconds  
(without alignments)  
2527.899 Million cell updates/sec

Title: US-09-762-767A-4

Sequence: 1 MYGFVNHAEELVIRNYGPE.....QVWFLSRKNTGTEETKQDDD 619

Scoring table: BLOSUM62

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Maximum DB seq length: 20000000000

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum	Match	100%
---------	-------	------

Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1	/SDS2/gcgdata/genseq/genseqp-emb1/AA1980.DAT.*
2	/SDS2/gcgdata/genseq/genseqp-emb1/AA1981.DAT.*
3	/SDS2/gcgdata/genseq/genseqp-emb1/AA1982.DAT.*
4	/SDS2/gcgdata/genseq/genseqp-emb1/AA1983.DAT.*
5	/SDS2/gcgdata/genseq/genseqp-emb1/AA1984.DAT.*
6	/SDS2/gcgdata/genseq/genseqp-emb1/AA1985.DAT.*
7	/SDS2/gcgdata/genseq/genseqp-emb1/AA1986.DAT.*
8	/SDS2/gcgdata/genseq/genseqp-emb1/AA1987.DAT.*
9	/SDS2/gcgdata/genseq/genseqp-emb1/AA1988.DAT.*
10	/SDS2/gcgdata/genseq/genseqp-emb1/AA1989.DAT.*
11	/SDS2/gcgdata/genseq/genseqp-emb1/AA1990.DAT.*
12	/SDS2/gcgdata/genseq/genseqp-emb1/AA1991.DAT.*
13	/SDS2/gcgdata/genseq/genseqp-emb1/AA1992.DAT.*
14	/SDS2/gcgdata/genseq/genseqp-emb1/AA1993.DAT.*
15	/SDS2/gcgdata/genseq/genseqp-emb1/AA1994.DAT.*
16	/SDS2/gcgdata/genseq/genseqp-emb1/AA1995.DAT.*
17	/SDS2/gcgdata/genseq/genseqp-emb1/AA1996.DAT.*
18	/SDS2/gcgdata/genseq/genseqp-emb1/AA1997.DAT.*
19	/SDS2/gcgdata/genseq/genseqp-emb1/AA1998.DAT.*
20	/SDS2/gcgdata/genseq/genseqp-emb1/AA1999.DAT.*
21	/SDS2/gcgdata/genseq/genseqp-emb1/AA2000.DAT.*
22	/SDS2/gcgdata/genseq/genseqp-emb1/AA2001.DAT.*
23	/SDS2/gcgdata/genseq/genseqp-emb1/AA2002.DAT.*

**Pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3231	100.0	619	21	AAV51608	Human soluble guan
2	1883	58.3	787	22	ABB58149	Drosophila melanog
3	929.5	28.8	690	22	ABB60360	Drosophila melanog
4	789.5	24.4	690	21	AAV70475	Human cyclic nucle
5	789.5	24.4	690	21	AAV51607	Human soluble guan
6	789.5	24.4	690	23	ABG61840	Prostate cancer-as
7	730.5	22.6	717	23	ABG61878	Prostate cancer-as
8	726	22.5	676	22	ABB58042	Drosophila melanog
9	704.5	21.8	669	22	ABB69903	Drosophila melanog
10	685	21.2	649	22	ABB65902	Drosophila melanog

[illegible]

## ALIGNMENTS

RESULT 1  
AAAY51608  
ID AAAY51608 standard; Protein; 619 AA

XX Human soluble guanylylcyclase beta1 protein.  
DE

KM Human,guanylylcyase alpha1; hscCalpha1; hscCbeta1; soluble;  
KM guanylylcyase beta1; antiarteriosclerotic; vasotropic; hypotensive;  
KM gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;  
KM peripheral arterial occlusive disease; arterial hypertension.

05 Homo sapiens

PN DE19837015-A1

PD 24-FEB-2000.

PF 14-AUG-1998; 98DE-1037015.

PR 14-AUG-1998; 98DE-1037015.

PA (VASO-) VASOPHARM BIOTECH GMBH & CO KG.

PI Schmidt H, Zabel U, Poller W;

DR WPI; 2000-184044/17

XX  
XX

PT encoding the subunits, useful for producing diagnostic antibodies, and  
 PT for somatic gene therapy of arteriosclerosis -

XX Example 1; Page 16-17; 44pp; German.

CC This invention describes novel purified human soluble guanylate cyclase  
 CC alpha1/beta1 (hgcalpha1/beta1). The products of the invention have  
 CC antiarteriosclerotic, vasotropic and hypotensive activity. Nucleic acid  
 CC sequences encoding the alpha1 and/or beta1 subunit are useful for somatic  
 CC gene therapy of arteriosclerosis and restenosis, ischemia (infarct),  
 CC peripheral arterial occlusive disease and arterial hypertension.  
 CC Antibodies to hgcalpha1/beta1 can be used for diagnosis of aberrant  
 CC hgcalpha1/beta1 expression in human tissues. This sequence represents  
 CC the human soluble guanylylcyclase beta1 subunit described in the method  
 CC of the invention.

XX Sequence 619 AA;

Query Match 100.0%; Score 3231; DB 21; Length 619;  
 Best Local Similarity 100.0%; Pred. No. 5,4e-306;  
 Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 1 MYGFVNHALLELVIRNTPGEVWEDIKKEAQLDEGGFLVRIYDSTKYDLVAASKVYN 60  
 DB 1 MYGFVNHALLELVIRNTPGEVWEDIKKEAQLDEGGFLVRIYDSTKYDLVAASKVYN 60  
 YY 61 LMAGEIIMQMGKMFVFCQSGSYDTILRVLSNVREFLQNDLADHDLATITPGMRAPSF 120  
 DB 61 LMAGEIIMQMGKMFVFCQSGSYDTILRVLSNVREFLQNDLADHDLATITPGMRAPSF 120  
 YY 121 RCTDAERKGLIILHYSEREGLODIVIGITKVAQIHGFEIDMKVYQORNECDHQFL 180  
 DB 121 RCTDAERKGLIILHYSEREGLODIVIGITKVAQIHGFEIDMKVYQORNECDHQFL 180  
 YY 181 IEERKEEDFYEDLDFFEENGTOESRIISPTFCAPPHIIPRDVLTQCGNAIRVL 240  
 DB 181 IEERKEEDFYEDLDFFEENGTOESRIISPTFCAPPHIIPRDVLTQCGNAIRVL 240  
 YY 241 POLQPGNCSLVSFSLVRPHIDISFHGILSHINTVFLRSKGLLDVEKLECEDELGTGE 300  
 DB 241 POLQPGNCSLVSFSLVRPHIDISFHGILSHINTVFLRSKGLLDVEKLECEDELGTGE 300  
 YY 301 ISCLRKGMIVYLPDASITFLCSPVNMDDLTRGLYLSDTPLHNTATDVLVLSGQFR 360  
 DB 301 ISCLRKGMIVYLPDASITFLCSPVNMDDLTRGLYLSDTPLHNTATDVLVLSGQFR 360  
 YY 361 EERKLOEELIDRLQTLRALDEKKKTDTLLYSVLPSVANELLHKKRPVPAKRYDNV 420  
 DB 361 EERKLOEELIDRLQTLRALDEKKKTDTLLYSVLPSVANELLHKKRPVPAKRYDNV 420  
 YY 421 TILFSGIVGFNAFCRSKASGEGAMKIVNLNDLYTRFDLTDSKRNPFYKVVETVGDXYM 480  
 DB 421 TILFSGIVGFNAFCRSKASGEGAMKIVNLNDLYTRFDLTDSKRNPFYKVVETVGDXYM 480  
 YY 481 TVSGLPSPCIHARSICHLALDMEIAGOVVDESVOITIGITGEVYVGVIGORMPRY 540  
 DB 481 TVSGLPSPCIHARSICHLALDMEIAGOVVDESVOITIGITGEVYVGVIGORMPRY 540  
 YY 541 CLFGNTVNLTSRTETGCKKINVSERYTCLMSPENDPOFHLHNGPVMKPKKPPMQ 600  
 DB 541 CLFGNTVNLTSRTETGCKKINVSERYTCLMSPENDPOFHLHNGPVMKPKKPPMQ 600  
 YY 601 VWFLSKRNTGTEETKODDD 619  
 DB 601 VWFLSKRNTGTEETKODDD 619

RESULT 2  
 ABB58149 ID ABB58149 standard; Protein; 787 AA.  
 XX AC ABB58149;  
 XX

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 1239.

KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO000171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW.

DR WPI: 2001-656860/75.

DR N-PSDB: ABL02252.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -

PS Disclosure; SEQ ID NO 1239; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 787 AA;

Query Match 58.3%; Score 1883; DB 22; Length 787;  
 Best Local Similarity 49.9%; Pred. No. 3.3e-174;  
 Matches 371; Conservative 102; Mismatches 125; Indels 146; Gaps 7;

YY 1 MYGFVNHALLELVIRNTPGEVWEDIKKEAQLDEGGFLVRIYDSTKYDLVAASKVYN 60  
 DB 1 MYGFVNHALLELVIRNTPGEVWEDIKKEAQLDEGGFLVRIYDSTKYDLVAASKVYN 60  
 YY 61 LMAGEIIMQMGKMFVFCQSGSYDTILRVLSNVREFLQNDLADHDLATITPGMRAPSF 120  
 DB 61 LMAGEIIMQMGKMFVFCQSGSYDTILRVLSNVREFLQNDLADHDLATITPGMRAPSF 120  
 YY 121 RCTDAERKGLIILHYSEREGLODIVIGITKVAQIHGFEIDMKVYQORNECDHQFL 180  
 DB 121 RCTDAERKGLIILHYSEREGLODIVIGITKVAQIHGFEIDMKVYQORNECDHQFL 180  
 YY 181 IEERKEEDFYEDLDFFEENGTOESRIISPTFCAPPHIIPRDVLTQCGNAIRVL 240  
 DB 181 IEERKEEDFYEDLDFFEENGTOESRIISPTFCAPPHIIPRDVLTQCGNAIRVL 240  
 YY 241 POLQPGNCSLVSFSLVRPHIDISFHGILSHINTVFLRSKGLLDVEKLECEDELGTGE 300  
 DB 241 POLQPGNCSLVSFSLVRPHIDISFHGILSHINTVFLRSKGLLDVEKLECEDELGTGE 300  
 YY 301 ISCLRKGMIVYLPDASITFLCSPVNMDDLTRGLYLSDTPLHNTATDVLVLSGQFR 360  
 DB 301 ISCLRKGMIVYLPDASITFLCSPVNMDDLTRGLYLSDTPLHNTATDVLVLSGQFR 360  
 YY 361 EERKLOEELIDRLQTLRALDEKKKTDTLLYSVLPSVANELLHKKRPVPAKRYDNV 420  
 DB 361 EERKLOEELIDRLQTLRALDEKKKTDTLLYSVLPSVANELLHKKRPVPAKRYDNV 420  
 YY 421 TILFSGIVGFNAFCRSKASGEGAMKIVNLNDLYTRFDLTDSKRNPFYKVVETVGDXYM 480  
 DB 421 TILFSGIVGFNAFCRSKASGEGAMKIVNLNDLYTRFDLTDSKRNPFYKVVETVGDXYM 480  
 YY 481 TVSGLPSPCIHARSICHLALDMEIAGOVVDESVOITIGITGEVYVGVIGORMPRY 540  
 DB 481 TVSGLPSPCIHARSICHLALDMEIAGOVVDESVOITIGITGEVYVGVIGORMPRY 540  
 YY 541 CLFGNTVNLTSRTETGCKKINVSERYTCLMSPENDPOFHLHNGPVMKPKKPPMQ 600  
 DB 541 CLFGNTVNLTSRTETGCKKINVSERYTCLMSPENDPOFHLHNGPVMKPKKPPMQ 600  
 YY 601 VWFLSKRNTGTEETKODDD 619  
 DB 601 VWFLSKRNTGTEETKODDD 619



KW cell proliferative disorder; autoimmune/inflammatory; diabetes mellitus;  
 KW neurological; vision; reproductive; smooth muscle.

OS Homo sapiens.

Key Location/Qualifiers

65 /note= "Potential phosphorylation site"  
 FT Modified-site  
 211 /note= "Potential phosphorylation site"  
 FT Modified-site  
 361 /note= "Potential phosphorylation site"  
 FT Modified-site  
 450 /note= "Potential phosphorylation site"  
 FT Modified-site  
 478 /note= "Potential phosphorylation site"  
 FT Modified-site  
 29 /note= "Potential phosphorylation site"  
 FT Modified-site  
 73 /note= "Potential phosphorylation site"  
 FT Modified-site  
 117 /note= "Potential phosphorylation site"  
 FT Modified-site  
 135 /note= "Potential phosphorylation site"  
 FT Modified-site  
 158 /note= "Potential phosphorylation site"  
 FT Modified-site  
 229 /note= "Potential phosphorylation site"  
 FT Modified-site  
 511 /note= "Potential phosphorylation site"  
 FT Modified-site  
 527 /note= "Potential phosphorylation site"  
 FT Modified-site  
 638 /note= "Potential phosphorylation site"  
 FT Modified-site  
 687 /note= "Potential phosphorylation site"  
 FT Modified-site  
 39 /note= "Potential phosphorylation site"  
 FT Modified-site  
 64 /note= "Potential phosphorylation site"  
 FT Modified-site  
 105 /note= "Potential phosphorylation site"  
 FT Modified-site  
 261 /note= "Potential phosphorylation site"  
 FT Modified-site  
 328 /note= "Potential phosphorylation site"  
 FT Modified-site  
 356 /note= "Potential phosphorylation site"  
 FT Modified-site  
 360 /note= "Potential phosphorylation site"  
 FT Modified-site  
 394 /note= "Potential phosphorylation site"  
 FT Modified-site  
 622 /note= "Potential phosphorylation site"  
 FT Modified-site  
 635 /note= "Potential phosphorylation site"  
 FT Modified-site  
 27 /note= "Potential phosphorylation site"  
 FT Modified-site  
 332 /note= "N-glycosylated"  
 FT Modified-site  
 479 /note= "N-glycosylated"  
 FT Modified-site  
 600 /note= "N-glycosylated"  
 FT Modified-site  
 472..661 /note= "N-glycosylated"  
 FT Modified-site  
 Region

/note= "Guanylate cyclase signature sequence  
 identified by PFM analysis: CNAP-3 shares 89%  
 identity with human soluble guanylate cyclase  
 large subunit."

W0200014248-A1.

16-MAR-2000.

XX 03-SEP-1999; 99WO-US20287.  
 PF  
 XX  
 PR 04-SEP-1998; 98US-0148904.  
 XX  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Hillman JL, Yue H, Guegler KJ, Corley NC, Patterson C, Tang YT;  
 XX WPI; 2000-256994/72.  
 DR N-PSDB; AAZ51684.  
 XX  
 XX Isolated cyclic nucleotide associated proteins useful for preventing,  
 PT diagnosing and treating cell proliferative, autoimmune/inflammatory,  
 FT neurological, vision, reproductive and smooth muscle disorders -  
 XX  
 PS Disclosure: Page 67-69; 78pp; English.

CC The present sequence is a human cyclic nucleotide  
 CC associated protein-3 (CNAP-3), identified in incyte clone 159278,  
 CC that is isolated from ADENINB01 cDNA library. It is expressed in  
 CC nervous, reproductive, cardiovascular and developmental tissues.  
 CC CNAP sequences may be used for prevention, treatment and diagnosis of  
 CC diseases associated with altered CNAP expression such as, cell  
 CC proliferative disorders (e.g. arteriosclerosis, cirrhosis, leukaemia,  
 CC lymphoma and cancer of the breast, prostate, lung and brain), autoimmune/  
 CC inflammatory disorders (e.g. asthma, anaemia, diabetes mellitus, multiple  
 CC sclerosis and psoriasis), neurological disorders (e.g. epilepsy, multiple  
 CC Alzheimer's/Parkinson's disease and strokes), vision disorders (e.g.  
 CC conjunctivitis, glaucoma, cataracts and retinitis pigmentosa),  
 CC reproductive disorders (e.g. infertility, uterine fibroids, ectopic  
 CC pregnancies and impotence) and smooth muscle disorders (e.g. angina,  
 CC anaphylactic shock, Kearns-Sayre syndrome and hypertension). The  
 CC coding sequence can be used for gene therapy.  
 CC  
 XX

Sequence 690 AA:

Query Match 24.4%; Score 789.5; DB 21; Length 690;  
 Best Local Similarity 35.1%; Pred. NO. 1.5e-67;  
 Matches 207; Conservative 96; Mismatches 219; Indels 67; Gaps 15;

QY 30 QLDEEGQFLVITITDDSKTYDVAASKVLMNGEIL-QMFGMFYFGQESYDILR 88  
 DB 101 KIKESKSLERDEDEKITAEOVAAGVP-----EVKESLGEVFRICTEED-ENILG 153  
 QY 89 VLGSNVEFLIONLDAL-----HDHIAITYPGMARSPFCCTDAERKGLILHYSERGLQ 143  
 DB 154 VVGGLTDLNFSFTLLKQSSHCQAGKRGLEDASTICDKR-DPLHNYFFPRKRTTS 212  
 QY 144 DIVIGIITKVAQQLHGFEDIMKVIQQRNECDHQ-----FLIEKESKEDFY 192  
 DB 213 LILPGIITKAANHVIYETEVEVSLM----PPCFHNDCEFPVNPYLLSVHKKSTKPS-- 265  
 QY 193 EDLRFRENGQESRISPYT-FCKAPPHIFEDLDLVVTOGNAIVYLPOLQ-PGNCSL 250  
 DB 266 -----LSPSKQSSLVIPSTSLFCTFPFHFFEDMDITLQFGKIRLMMRRDPQGPNF 320  
 QY 251 LSVESLVRPHIDISFHGILSHINTVFYLRKSGLLDVEKLECEDELGTETISCLRLGOM 310  
 DB 321 EEFEEILTPKINQFSGIMTLMNQFVVRBMNSVKR-----SSRVMDLKGOM 370  
 QY 311 IYLPEDASILFLCSPSVNDDLLRRLGLYSDIPLHATRDVYLGGQFPEEYLTQELE 370  
 DB 371 IYVESAILFLGSPCVDRLEDFTRGLYSDIPRIHNAIRDVYVIGBARADQLKRLG 430  
 QY 371 ILTRDLRLALDEDEKKTDTLLYSVLPVANELHKKRPVPAKRDNTYITSGVGF 430  
 DB 431 KLRATLQADQALDEEKKKYDLCSTPCEVAQQLQOGVOYVKKSTNTMFSIDVGF 490  
 QY 431 NAFCSKHAEGGAKRIYNLLNDLYTRFDLTDSRKNPFFYKVEIVGDKRYMTVSGILPPCI 490  
 DB 491 TAISSQCS-----PLOYITMLNALTTRFDQCGELD---YKVEIIGAYCAYAGLHRES 543



```

OY 491 HHAASICHIALDMMETAGOV-QVDGESVOITIGITGEVYVIGCGMRPRTCLFGNTVL 549
DB 544 THAVOIALMALKMEIISDEVMSPHGEPIKMRIGLSGVSFAGVGVKMPRYCLFGNNVTL 603
OY 550 TSRTETGKRGKINSEYTRCL-----MSPENDPOFHLEHRS 588
DB 604 ANKFESCVPRKINVSPTTYRLKDCPGFVTPPRSREELPPNPSEIPG 652

RESULT 5
AAY51607
ID AAY51607 standard; protein; 690 AA.
XX
AC AAY51607;
XX
DE 26-MAY-2000 (first entry)
XX
DE Human soluble guanylylcyclase alpha protein.
XX
KW Human; guanylylcyclase alpha; hscGalpha1; hscGbeta1; soluble;
KW guanylylcyclase beta1; antiarteriosclerotic; vasotropic; hypotensive;
KW gene therapy; arteriosclerosis; restenosis; ischemia; diagnosis;
KW peripheral arterial occlusive disease; arterial hypertension.
XX
OS Homo sapiens.
XX
PN DEL9837015-A1.
XX
PD 24-FEB-2000.
XX
PF 14-AUG-1998; 98DE-1037015.
XX
PR 14-AUG-1998; 98DE-1037015.
XX
PA (VASO-) VASOPHARM BIOTECH GMBH & CO KG.
XX
PI Schmidt H, Zabel U, Poller W;
XX
DR WPI: 2000-184044/17.
XX
DR N-PSDB; AAZ88938.
XX
PT New human soluble guanylate cyclase alpha/beta and the nucleic acid
PT encoding the subunits, useful for producing diagnostic antibodies, and
PT for somatic gene therapy of arteriosclerosis -
XX
XX
XX Example 1; Page 13-15; 44pp; German.
XX
CC This invention describes novel purified human soluble guanylate cyclase
CC alpha/beta1 (hscGalpha1/beta1). The products of the invention have
CC antiarteriosclerotic, vasotropic and hypotensive activity. Nucleic acid
CC sequences encoding the alpha and/or beta subunit are useful for somatic
CC gene therapy of arteriosclerosis and restenosis, ischemia (infarct),
CC peripheral arterial occlusive disease and arterial hypertension.
CC Antibodies to hscGalpha1/beta1 can be used for diagnosis of aberrant
CC hscGalpha1/beta1 expression in human tissues. This sequence represents
CC the human soluble guanylylcyclase alpha subunit described in the method
CC of the invention.
XX
XX
XX Sequence 690 AA;
XX
OY Query Match 24.4%; Score 789.5; DB 21; Length 690;
OY Best Local Similarity 35.1%; Pred. No. 1.5e-67;
OY Matches 207; Conservative 96; Mismatches 219; Indels 67; Gaps 15;
DB 30 QLDGEGFLVRIIVYDSKTYDVAASRVNLNAGRII-QMFGKMFVFCQESGYDTIR 88
DB 101 KIKESKRSIERDEFEKTIAEOVAAGVP-----EVIESIGEEVFKICYED-ENILG 153
OY 89 VLSNVREFLQNLDAI-----HDHLATIVPGKAPSFCTDAEKGGKGLIHVYSEREGIQ 143
DB 154 VVGGLKDFLNFSTLKGSSHQEGKRGRLDASIIICLDE-DFLHVYVFFPRRTS 212
OY 144 DIVIGIKTVAOOIHGTETIDMKVIOQRNECHTO-----FLIEKESKDEDFY 192

```

```

DB 213 LILPGIKKAHVAVLYETEVEVSLM-----PCFHNDCSEFVNOPYLLYSVHMSTRPS--- 265
OY 193 EDLDREENGTOESRISPT-FCKAPEFHIIIDRDLYVQGNATYRVLPQ-IGNCSL 250
DB 266 -----LSPSKPOSSILPTSLFCKTFPFHMFMDKMTILQFNGIRLRNRDFOCKPWF 320
OY 251 LSEFSLVRPHIDISFNGILSHINTVFLRSKGLDVERKECEDELTGEISCLRKGOM 310
DB 321 EEFPEILFPKINOTFSGIMTLMNQFVAVRRRDNVSKR-----SRVMDLKGOM 370
OY 311 IYLPREADSLFLCSPSVNMLDITFRRLGLYSDIPLHDATRDLYLGEQFREERYKLTQLE 370
DB 371 IYVESALLFLFSGPCVDRLEDFTGRLYSIPINHALRDVYLIGEQAADGLKRRIG 430
OY 371 ILTDRLQTLRALDEKKTDTLLYSVLRPSVANEIRHRPAPKRYDVTILFSGIYVF 430
DB 431 KIKATLEQAHQALDEKKTVDLCSIFCEVAQQLMOGQVQAKKFSVNTLPSDIVGF 490
OY 431 NAFCSKHASGEGAMKIVNLNDLYTRFDLTDSRKNPFYKRYETVGDKYMTVSGLPDPCI 490
DB 491 TAIQSQCS-----PIQYITMLNLTYRFDQCGELD---YKYETIGDACYVAGGLHKESD 543
OY 491 HHAASICHIALDMMETAGOV-QVDGESVOITIGITGEVYVIGCGMRPRTCLFGNTVL 549
DB 544 THAVOIALMALKMEIISDEVMSPHGEPIKMRIGLSGVSFAGVGVKMPRYCLFGNNVTL 603
OY 550 TSRTETGKRGKINSEYTRCL-----MSPENDPOFHLEHRS 588
DB 604 ANKFESCVPRKINVSPTTYRLKDCPGFVTPPRSREELPPNPSEIPG 652

RESULT 6
ABG61940
ID ABG61940 standard; protein; 690 AA.
XX
AC ABG61940;
XX
DE 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #141.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytoslastic.
XX
XX
XX Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US32045.
XX
PR 13-OCT-2000; 2000US-0687576.
XX
PR 08-DEC-2000; 2000US-0733288.
XX
PR 08-DEC-2000; 2000US-0733742.
XX
PR 24-JAN-2001; 2001US-263957P.
XX
PR 16-MAR-2001; 2001US-276791P.
XX
PR 06-APR-2001; 2001US-276888P.
XX
PR 24-APR-2001; 2001US-286214P.
XX
PR 30-APR-2001; 2001US-0847046.
XX
PR 04-MAY-2001; 2001US-288589P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afari D, Hevezl P;
XX
DR WPI: 2002-471335/50.
XX
DR N-PSDB; ABK92259.
XX
PT detecting a prostate cancer-associated transcript in a cell in a
PT patient; useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes

```



DB 294 MTILPFGNGIRRLAMRROFQCKPNEFEYFEILLPKINQTFSGIMTMNQFVVRVRMRDMS 353  
 QY 287 VEKLECEDELGTETISCLTKCOMITYPEADSIILFCSPSVANLDDLTRRGHLYSDIPLH 346  
 DB 354 VKK-----SSRVMDLKCOMIYIVESSAILFLGSCVDRLEDFTRGHLXSLDIPRH 403  
 QY 347 DATRLVLLGEOFREERYKLTQELTLRLQLTLALDEDEKKTPTLLYSVLPSPVANL 406  
 DB 404 NALRQVVLIGEQAADQGLKRLKATLEQAHOALEEKKKTVLLCSIFPCEVAQOL 463  
 QY 407 RHKRPVPAKRDYNTILFSGIYGFNAFCSKHSAGSAGIVLNDLYTRPTLDSRN 466  
 DB 464 WOGVQVQAKFSNVTMLFSDIYGFALCSQCS---PLQVITMLNLYTRFDQCGEID- 518  
 QY 467 PFVYVETVGDKYMTVSGILPEPCIHARSICHLALDMEIAGV--QVDESVQITIGIT 525  
 DB 519 --VYVETIAMPITVILGSLHKSDFTHAVQIALMKLMELSDVMSPHPEPIKMRGLHS 576  
 QY 526 GEVTVGVIGQMRPRICGNTVNLASRTETTGCKINSEYTYRCL-----MSP 575  
 DB 577 GSVFAGVYGVKMPRCLEGNVTLANKESSCVPRKINVSPTTYRLKDCPGFVETPRSR 636  
 QY 576 ENSDPOFHEHRG 588  
 DB 637 EELPFPFSEIRG 649

RESULT 8  
ABB58042

ID ABB58042 standard; Protein; 676 AA.

XX ABB58042;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 918.

KW Drosophila: developmental biology; cell signalling; insecticide;  
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

XX WPI, 2001-656860/75.

XX N-PSDB; ABL02145.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signalling and cell-cell  
interactions -

XX Disclosure: SEQ ID NO 918; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABB5737-ABB72072) and the encoded proteins  
XX (ABB5737-ABB72072).  
XX The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIP0  
at ftp.wipo.int/pub/published\_pot\_sequences.

XX Sequence 676 AA:

Query Match 22.5%; Score 726; DB 22; Length 676;  
Best local Similarity 30.4%; Pred. No. 2.3e-61;  
Matches 187; Conservative 125; Mismatches 235; Indels 68; Gaps 18;

QY 22 WEDIRKQALDEGGQVLRIIYDSTKYDLVAASVYL-----NLNAGELIQFGKRFYF 77  
 DB 73 WPNTRK-LKLDPO-----TFRCANVYLYLDIQELLMKMDASASEILLVGLBELTIC 124  
 QY 78 COESGYDILRYLGSNVREFIQNLDAHLHLATTYGMRAPSRCRDARKGKGLIHYYS 137  
 DB 125 CCGIIEIRAFRCIGTDLGFLGSLGSDGVYLNKQEDVDYDTGVCA-----GEBELI-FTS 179  
 QY 138 EREGADIDYIGIKTVAAQOIHGETIDMKYIQORNE-----CDHTQ-FLIEEKES 186  
 DB 180 ERVVIAMLLIGSLIKALTRMLYKVDVNIKIEPEGDARRYRLFSLVKDNSQFTLMGRPTS 239  
 QY 187 KEEDFEYEDLDREENGTOESRISPYFECAAPPHILFDDDLVYTOGGMATYRL-QQLP 245  
 DB 240 VSKTIPETVQORSNSNASDQNMSSFCMFPWHFTMNOQLLEVQLGRGFSKLYKPYMD 299  
 QY 246 GNCSTLSVPSLVPRH-IDISFHLISHWVYV--LRKRGLLDYKLECEDELGTETIS 302  
 DB 300 FGQQAITYFDKRPKGLTKKFRDIYARRTYTPFLIGLNNPGVANDPFAIG----- 348  
 QY 303 CLRLKQOMITYPEADSIILFCSPSVANLDDLTRRGHLYSDIPLHADRDLVLGEOFRE 362  
 DB 349 -LEIKQAVHCESNSILFISGPFILDLGICNGIFISDIPLDATREVIIVGEQARAQ 407  
 QY 363 YKLTQELTLDRLQLTLALDEDEKKTPTLLYSVLPSPVANLRRKRPVPAKRYNTI 422  
 DB 408 DGLRRMDKIKNSIEEANSVTERKKNVSLHLFPAETAEKLMGSSIDAATYDVTI 467  
 QY 423 LFSGIFNAFCSKHSAGSAGAMKIVLNDLYTRPTLDSRNKP--YKYVETVGDKYM 480  
 DB 468 LFSDIYGFISCSRAI--PFVYISMLGSLVKRDFEPCD-----FFDYKNETIDDAIC 518  
 QY 481 TVSGILPEPCIHARSICHLALDMEIAGV--VOVDESVQITIGITGEVTVGVIGQMRP 539  
 DB 519 VASGLIRASIDYAHKAVAMMALMKMIDACSKHITHDGRQIKRIGLHGTIVLAGVGRKMP 578  
 QY 540 YCLFGNTVLTSTRTETTGCKINSEYTYRCLMSPENDPQFHLHRPVSMMKGKKEPM 599  
 DB 579 YCLFGHSVTIANKFESGSEALKINVSPTTKDWLTKEGF--EPELQRPD-SFLPRFEP- 634  
 QY 600 QWFLSRKNTGTEET 614  
 DB 635 -----NPGGTET 641

RESULT 9  
ABB69903

ID ABB69903 standard; Protein; 669 AA.

XX ABB69903;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 36501.

KW Drosophila: developmental biology; cell signalling; insecticide;  
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

Pf	23-MAR-2001; 2001WO-US09231.
Xx	
Pr	23-MAR-2000; 2000US-191637P.
Pr	11-JUL-2000; 2000US-061415O.
Xx	
Pa	(PEKE ) PE CORP NY.
Xx	
Pt	Venter JC, Adams M, Li PWD, Myers EW;
Dr	WPI: 2001-656860/75.
Nr	N-Psdb; ABLI4006.
Xx	
Pt	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
Xx	
Ps	Disclosure: SEQ ID NO 36501: 21pp + Sequence Listing; English.
Xx	
Cc	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLI6176-ABL3511), expressed DNA sequences (ABLI01840-ABLI6175) and the encoded proteins (ABBS7737-ABBS2012).
Cc	
Cc	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
Xx	
Sq	Sequence 669 AA:
	Query Match 21.8%; Score 704.5; DB 22; Length 669;
	Best Local Similarity 28.8%; Pred. No. 2.9e-59;
	Matches 197; Conservative 123; Mismatches 264; Indels 101; Gaps 18;
Oy	1 MYGFNHALTELLVIRKNPGVEWMEDIKRPAODE--EGQFIVRIIYDSDSKTDLVAAASKV 58.
Db	1 MYGMATYESVOHITIQOEYEMETM---RKVCQIYDCKHOSFKTHQIIIPDKIMPDPFAA---- 53
Oy	59 LNLNAGE---LLOWGKMFVFECQESGYDTILRVLGSNVEEFLONLDALHDHLATIYPG 114
Db	54 LSASTGESFDCECMNFEGRCFVRFFSNFGYDMKINRGTFCDQLSIDINHQMFTYPK 113
Oy	115 MRASFRCTDAEKGGGLIHYSSEEGODIYIGIKIVVAOIHTETIDMKYIOQRNEC 174
Db	114 MKSSSMOLTTAND-DGAVILYRSGRGTSGSKYLIGOMTVNAEEFGDLMTAYVLESQNDIC 172
Oy	175 DHFOFLLEEKSKDEDFEDRDREENTQESRIPLY-----TFCAPFPH 220
Db	173 GGTAGPIRLTEGPLVIYKYRLDDNRMYAKRVAVIANHPQLMKPSVDLANVLELFPT 232
Oy	221 IIFPDRLVYTCCGNATRYVLPOLOG-----NC----- 248
Db	223 IVLDHDMKITLAGERIVETWLINHPGVNPKFTGISHLIERKCRRPKTOLOMETILOMR 292
Oy	249 SLIASVESLYR-----PHIDSFHGI-----LSHINTFYLRKSREGILDYELEC-- 292
Db	293 TVLEFEELIRKGHNRAADALANDEFENFDAASSLNQAAMALASAKFSSENAREEA 352
Oy	293 ----EDEL--TGTE-----ISCLRLKQMITLPBADSILFCSPSVNLDDLTPRGLYL 340
Db	341 SDIFLMDTRQVLLGBOFREERYKT-QELELDRQLTLRALEDKKRTDLLSYLP 399
Oy	413 NDLPNHGLSRELVMAGOWCHSKLEIMFEKEORSDLEKSLSELADSWKRQGBELLYSMP 472
Db	400 PSVANLELHKRPVPYAKRKYNTLTISGIVGNAFSCASAEGAKIYNLINDLTYTRDJ 459
Oy	473 RPIAEPRMKSEHNQCQSFEELYVIEFM--NIYDSGSNNIYODMANQATYTLTKV---TSA 527
Db	460 LTDSRKNDPYVKVETVGDKTYVTSGLPKCIHARSICHLLADMELTAGOVQVDESVOI 519

```

Db      528  LDEELISFFVYKVEYGVNVAASGAPVYNLHAHACDLALRWKKAKAHALPG--VAI 585
Oy      520  TIGHTHTGVTYGVIGQRRPRYCLGNTVNLTSRTETTGKKKINVSERYTCLMSPEND 579
Db      586  RVGINSNGVYAGVYGMKVPKRYCLFGDVTNTASRMSSSDPMWIOISNTY-----ALKVOK 640
Oy      580  PQFHLHRCGPSMKGRKEPMQWFL 604
Db      641  VGKVEARGFVYKKGKGMETTYWLL 665

RESULT 10
ABB69902
ID      ABB69902 standard; Protein: 649 AA.
XX
XX      ABB69902;
AC
XX      26-MAR-2002 (first entry)
DT
XX      Drosophila melanogaster polypeptide SEQ ID NO 36498.
DE
XX      Drosophila; developmental biology; cell signalling; insecticide;
KW      pharmaceutical.
XX
XX      Drosophila melanogaster.
OS
XX      WO200171042-A2.
XX      27-SEP-2001.
XX
XX      23-MAR-2001; 2001WO-US09231.
XX
XX      23-MAR-2000; 2000US-191637P.
XX      11-JUL-2000; 2000US-0614150.
XX
XX      (PEKE ) PE CORP NY.
XX
XX      Venter JC, Adams M, Li PWD, Myers EW;
XX      WPI; 2001-656860/75.
XX      DR      N-PSDB; ABL14005.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
XX      PT
XX      PS      Disclosure; SEQ ID NO 36498; 21pp + Sequence Listing; English.
XX
XX      The invention relates to an isolated nucleic acid detection reagent
XX      capable of detecting 1000 or more genes from Drosophila. The invention is
XX      useful in developmental biology and in elucidating cell signalling and
XX      cell-cell interactions in higher eukaryotes for the development of
XX      insecticides, therapeutics and pharmaceutical drugs. The invention
XX      discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX      sequences (AB101840-AB16175) and the encoded proteins
XX      (ABB57737-ABB72072).
XX      The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pct_sequences.
XX
XX      Sequence 649 AA;

Query Match      21.2%; Score 685; DB 22; Length 649;
Best Local Similarity 28.5%; Pred. No. 2.2e-57;
Matches 193; Conservative 122; Mismatches 256; Indels 106; Gaps 18;

Oy      1  MYGFVNALAEILVIRNYGPEVWEDIKKEAQDEGQFVIRIITYDSKTYDLVAASKVLN 60
Db      1  MYGMLYESVQHYVGEYGVDMRWKVCIIIDC-KINSFRTHTIYDKLMPDLAEALISACTG 59
Oy      61  LMGCELLMGKMFVFCQESGYDTILRVIGSNVREFLNIDALDHLDLATITPCMRAPSF 120

```

```

Db DN-----FGYDKMIRSTGTYFCDFLOSIDNHLNMFYPMKSPM 101
QY 121 RCTDAEKGGLILAYSEREGLODVIIGIKTVAQOIHGETIDMKVIOQNECDHT--- 177
Db 102 QLTNMD-DNGAVIILYRSTRGSMKYLIGOMTEVAREFGLIEKAVIYESQNDISGTAAP 160
QY 178 -----QDLIEKEKEDEDFEYEDLRFENGQOEIRISYT-----FCKAFPHIIFDRD 226
Db 161 IKLDGPIPLTVIKKRLDFDNREYMAKRVNTEAHPQSOLKMPYKLDVFLDFLFPFTFLVLMND 220
QY 227 LVTNOCGNATYRVLPLQDPC-----NC-----SLLSYF 254
Db 221 MKITHAGKIVETVIMHNPANPKSFIGHVWDLQCRRPKDTTIDMTLIMRAVLPEF 280
QY 255 SLVPR-HIDISFHT-----LSHNTFVLSRKE-----GLDVEKLECEDEL- 296
Db 281 ELIRFGHNRRAAYDAVLNMFENYDEMDINEAQTALAKQFSESHPVDDDESARDEID 340
QY 297 --TGTE-----ISCLRLKGOMTYLPEADSIILFCSPSVANLDDLRRGILYLSIDIPLHAT 349
Db 341 PATGERRSSQGLRSTLLKGOMFYINDVSLIFLCSPLIENLDELHGIGLYLNDLNPHGIS 400
QY 350 RDVLALGEOFREERYKLT-QELEILTRLOLTLRALEDEKKTKDTLLYSVLPSPVANELRH 408
Db 401 RELVMAAGQHCCKLEIMEKEKORDELEKSLADSKRQDELLYSMPPIARML 460
QY 409 KRVPAPKRYDNVTILFSGIYGFNAFCSKHAGSEGAAMKIVNLNLTREDTLLTRSKNPF 468
Db 461 SEEGQCSGFEEYSVIFLEVM--NVYDEGLNSIGAMQVNTLNKV---FSALDEELISPF 515
QY 469 YKVTVDGDKYTVASGLPEPCIHARSICHLADMMELAGOVQV-DGESVOTTIGIHCE 527
Db 516 YKVTVMGVYMAVAGADVPNPLAHACDLALRYMK---KRAHMDGDVALRVGINSGP 572
QY 528 VVTGIGORMPRYCLFGNTVNLTSRTETGEGKINVSERYRCILMSPENSDPOFLEHR 587
Db 573 VVAGVVGOKVRPRYCLFGDTVNTASRMSESSDPKTIQLSKYT-----GDKVROGVKVESR 627
QY 588 GPVSMKGRKEPMQVWFL 604
Db 628 GTVQYKRGKDMETWLL 644

RESULT 11
AA00754
ID AAC00754 standard; Protein; 119 AA.
XX
AC AAC00754;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 4835.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PS N-PSDB; AAC00760.

```

```

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
PS Claim 13; SEQ ID 4835; 71pp + CD-ROM; English.
XX
CC The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3',
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
SQ Sequence 119 AA;
XX
Query Match 18.5%; Score 599; DB 21; Length 119;
Best Local Similarity 98.3%; Pred. No. 3.3e-50;
Matches 117; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MYGFVNAHELVIRNYGPEVWEDIKKEAQDDEGQFLVRIITDYSKTYDLVAASKVYN 60
Db 1 MYGFVNAHELVIRNYGPEVWEDIKKEAQDDEGQFLVRIITDYSKTYDLVAASKVYN 60
QY 61 LNAGETLQMGKMFYFCOESGYDTLLRVLSNVREFLONDLADHDLATYIPGMKAPS 119
Db 61 LNAGETLQMGKMFYFCOESGYDTLLRVLSNVREFLONDLADHDLATYIPGMKAPS 119

RESULT 12
AA011280
ID AA011280 standard; Protein; 1061 AA.
XX
AC AA011280;
XX
DT 12-MAR-2002 (first entry)
XX
DE Human atrionatriuretic peptide receptor A (NPR1) protein.
XX
KW Human; natriuretic peptide receptor A/guanylate cyclase A; NPR1;
KW atrionatriuretic peptide receptor A; haplotyping; cytostatic; genotyping;
KW haplotype pair; single nucleotide polymorphism; gene therapy;
KW drug screening; hypertension; hypotensive.
XX
OS Homo sapiens.
XX
PN WO200179231-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US12300.
XX
PR 14-APR-2000; 2000US-197330P.
XX
PA (GENA-) GENAISSANCE PHARM INC.
XX
PI Bentivegna SC, Choi JY, Kiem SE, Nandabalan K;
XX
DR WPI; 2002-066340/09.
XX
PS N-PSDB; AAS16995.
XX
PT Genotyping human natriuretic peptide receptor A/guanylate cyclase gene
PT of an individual, involves determining identity of nucleotide pair at
PT specific polymorphic sites for two copies of the gene
PS Claim 27; Fig 3; 96pp; English.

```



Sequence 1061 AA;

Query Match 14.9%; Score 481.5; DB 23; Length 1061;  
 Best Local Similarity 27.7%; Pred. No. 3,6e-37;  
 Matches 197; Conservative 88; Mismatches 227; Indels 199; Gaps 31;

33 EGGQFLVRIYDDSTKYDYLAASKVLN-----LNAGET 66  
 416 ENGARVAVLNTNGT-SQELVAVSGKRLNMPLOYPPEDIPKCGFEDNEDPACNODHSTLEV 474  
 67 LQMFQKM-----CFEVE-----COESGYDILRYLGSNVEFLQNALDHLATIVP 113  
 475 LALVGSLSLGLIIVSFYIRKMOLEKELASLMKRWEDVDP-----SLERHL----- 524  
 114 GNRAPSFCTDAEKK--GLLIHYYSERGLQDIYIGIKTYAQOIHGTEDIMKYIQOIN 171  
 525 --RSAGSRLLTSGRSGNSGSL-----TTGQFOV--FAKTAAYK--GMLVAVKRVNRK 573  
 172 EECDDTQPLIEKESKEEDFYEDLDREFENGTOESKISPT--FCAAFPHIIFDRLVVT 230  
 574 IEL--TRKVFELKMRDQNHRLRFVGACTDPNLCILTEYCPGSLQDILEMSITL 631  
 231 -----QCGNATRYVLPOLQPGN--CSLSVFLVRPHIDISPHGISHINTV---FVLR 279  
 632 DMKFRYSLENDIVKGMFLHNGAICS-----HGNLKSSNCVVDGFRVLK 675  
 280 SKE--GLLDVEKLECEDELGTETISCLRLKGMIV---LPEADSLFLCSPSV--NNLDDL 333  
 676 IDYGLSEFRDLDP-----QGHVYAKKLMTAPELLMAASPPVRSQAGDV 722  
 334 TRRGVLSIDIP-----HATRD-----VLGEO--FREKYLQDELEL----- 372  
 733 YSFGILDEIALRSGVFHVEGLDLSPKETIERTNGEOPFRSIALQSHLEELGLMOR 782  
 373 -----TDRLQTLR-----ALEDER 387  
 783 CMAEDPQERPPQRLTIRKRNRENSNIIDNLSRMQYANNLELYEERQATLEEK 842  
 388 KKTDTLLYSVLPSPVANELRHRPVPKRYDNTVILFSIGVFNAFSCRHASGEMKIV 447  
 843 RAEALVLYOILPHSAEQLEKRETVQAEAFDSTVIYFSDIVGTALSAEST---PMQV 898  
 448 NLNLNLYRFDLTLTSRKMPFYKVTYGDXYMTVSGLEPCH---TARSCHLALDM 503  
 899 TLNLDLYTCEDAVID---NEDYKVTETIDATNVVSGP---VRNGLRACEARALAL 952  
 504 MEIAGOVQY---DGEVQITIGIHGEVNTYIGVGMPPRYCLFGANTVNLTSRTETGEGK 560  
 953 LDVAVSFRIHRPQDLRLRIGIHQPCVAGVGLKMPRYCLFGDVTAVNATSRMESGEL 1012  
 561 KINVESEYTYRCIMSPENDPQHLHRGVPVSMKKKEPMQVFLSKRNTGT 611  
 1013 KIHLSSETKAVL-----EEFGFELKLDGVEMKKGKVTYMLGERSST 1059

RESULT 14  
 ABB11783  
 ID ABB11783 standard; peptide; 1075 AA.  
 XX  
 AC ABB11783;  
 XX  
 DT 11-JAN-2002 (first entry)  
 XX  
 XX Human ANP-A receptor homologue, SEQ ID NO:2153.  
 DE Human ANP-A receptor homologue, SEQ ID NO:2153.  
 KW Human: cytokine; cell proliferation; cell differentiation; growth factor;  
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;  
 KW atherosclerosis; coronary heart disease; arterial ischaemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;

KW tissue regeneration; wound healing; infection; immune disorder;  
 KW cell culture; drug screening; gene therapy; antiinflammatory;  
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
 KW antifungal; vulnereary; antulcer.  
 OS Homo sapiens.  
 PN WO200157188-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US03800.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 XX  
 PA (HYSR-) HYSRQ INC.  
 PI Tang YT, Liu C, Drmanac RT;  
 DR WPI; 2001-457740/49.  
 XX  
 DR N-PSDB; ABA09027.  
 PT Human proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PS e.g. arthritis and cancer -  
 XX  
 PS Claim 20; Page 246-247; 1963pp; English.

Sequences ABB10981-ABB1230 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth factor activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a novel human polypeptide of the invention.

Sequence 1075 AA;

Query Match 14.9%; Score 481.5; DB 22; Length 1075;  
 Best Local Similarity 27.7%; Pred. No. 3,7e-37;  
 Matches 197; Conservative 88; Mismatches 227; Indels 199; Gaps 31;



```

QY 33 EGGQFLVRIYDSDKYVLAASAKVLA-----LNAGEI 66
D 430 ENGAFRVLYNNGT-SQELVAVSGRLNMPGLGPPDPICGFDNEDPACNOHLSTLEY 488
QY 67 LQMFQK-----FFVF-----COESYDILRVGSVNRELOLDLADHLATYIP 113
D 489 LALVGSLSLGLILVSPFYRKMOLEKELASELMRWEEVER-----SSLERL----- 538
QY 114 GMRAPSRCTDAEKGK-GLILHYSSREGIADIVIGITKVAQOIHGETIDMKVIOQRN 171
D 539 --RSAGSRLLSGSGSYGSL-----TTGQFOV---FATATYK--GNLYAKRVKRR 587
QY 172 EECDFQFLIEEKSEKEDFYEDLDREENGQDESRIISPYT-FCKAPFHILFDRLVVT 230
D 588 IEL--TRKVLFEKHMVDVONENHILTRFVGACTDPNCCIITEXCPRGSLADILENESITL 645
QY 231 -----QCGNAIRVLPQLOPGN-GLLSVFSVLRPHIDISFNGILSHIMTV-----VYLR 279
D 646 DMFRRSLNDIVKGMFLHNGAICS-----HGNLSSMCVVDGRVYK 689
QY 280 SKE-GLLDVKELECEDELGTETISCLRLKGMITY---LPEADSTILFLCSFSV--MNLDDL 333
D 690 ITDYGLESFRLDPE-----QCHTYAKKLMTAPELLNRASPPVRSQAGDV 736
QY 334 TRRGVYISDIPL-----HDATRDL-----VLGEO--FREBYKLTQLELIL----- 372
D 737 YSPGILLOETALRSQVGFHVHGLDLSPEKILIERVTYRGEPFRSLALQSHLEELGMOR 796
QY 373 -----TDRIQLTR-----ALDERK 387
D 797 CAEDPOERPPFOQIRLTLRKFNRENSNILLNLSMEOYANNLEELVEBRTQAYLEEK 856
QY 388 KKTDTLLYSVLPSPVANELHKKRPVPAKRYDNTILFSGIVGFNAFCSKHAEGAKIY 447
D 857 RKAEALLYQLPHSVAQLRGETVQAEDSVITYSDIVGTALSAEST---PMQVY 912
QY 448 NLNDLYTRDITLDSKKNPFYKVEYQKRYMVSGLPEPCIH---HARSICHLALDM 503
D 913 TLLNDLYTCFDAVID---NEDVYKVERTIGDAYVVSGLP---VANGRLHACEVAFMALAL 966
QY 504 MEIAGOVV---DQESVQITIGITHGEVYGVIGORBPYCLFGNTVNLASRPTTEGK 560
D 967 LDVRSRIRHRRQEQRLRLKIGITHGVCAGVGLKMPRYCLFGDVTYNASRRESNGEAL 1026
QY 561 KINVSERYTYCLMSPENDPOFHLHRGPVSMKGGKPEMQWFLSRKNTGT 611
D 1027 KIHLSSETKAVL-----EEFGFELRLGVDYEMKKGKVRITYWLLGERSST 1073

```

```

XX (UYVA-) UNIV VANDERBILT.
XX Garbers DL, Schulz S;
XX WPI; 1993-272183/34.
XX New purified enterotoxin receptor protein - used to develop
XX prods. for treating abnormal conditions caused by bacterially
XX released enterotoxin, partic. diarrhoea
XX Disclosure; Fig 3; 26pp; English.
XX The sequences given in AAR38862-63 represent the guanylyl cyclases,
XX GC-A and GC-B, which binds heat stable enterotoxin. These proteins
XX are enterotoxin receptors which may be used as a therapeutic to control
XX intestinal fluid permeation as well as abnormal conditions caused
XX by bacterially released enterotoxin. The binding domain of the
XX proteins, or antibodies to the proteins, can be used to eliminate
XX diarrhoea. The proteins may be used to isolate ligands and to screen
XX for antagonists of toxin binding.
SQ sequence 1029 AA:
Query Match 14.58; Score 468; DB 14; Length 1029;
Best Local Similarity 43.48; Pred. No. 7.2e-36;
Matches 112; Conservative 38; Mismatches 82; Indels 26; Gaps 9;
QY 361 EYKLPQELILNDRLQTLRALEDEKKKDTLLYSVLPSPVANELHKKRPVPAKRYDNY 420
D 789 EYANMLE-ELVEBRTQAYL-----EKKRAEALLYQLPHSVAQLRGETVQAEDSV 843
QY 421 TILFSGIVGFNAFCSKHAEGAKIYVNLNDLYTRDITLDSKKNPFYKVEYVGTGDKYM 480
D 844 TIFYSDIVGTALSAEST---PMQVYVTLNDLYTCFDAVID---NEDVYKVERTIGDAYM 896
QY 481 TVSGLPPEPCIH---HARSICHLALDMETAGOVV---DQESVQITIGITHGEVYVTV 533
D 897 VVSGLP---VRNGQLHAREVAFARMAALLDVRSFRIRHRRQEQRLRLKIGITHGVCAGV 953
QY 534 GORMPRYCLFGNTVNLASRPTTEGKINVSERYTYCLMSPENDPOFHLHRGPVSMK 593
D 954 GLKMPRYCLFGDVTYNASRRESNGEALKIHLSSETKAVL---EEFD-GFLELRGVDYEMK 1009
QY 594 GKKEPMQWFLSRKNTGT 611
D 1010 GKGVRYTYWLLGERSST 1027

```

Search completed: June 27, 2003, 13:03:03  
Job time : 35.6287 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:03:07 ; Search time 17.4966 Seconds  
(without alignments)  
3401.080 Million cell updates/sec

Title: US-09-762-767a-4

Perfect score: 3231

Sequence: 1 MGFVNHAELELVIRNYGPE.....QWFLSRKNTGTEETKQDD 619

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3231	100.0	619	1 S23097	guanylate cyclase
2	3206	99.2	619	1 OYB070	guanylate cyclase
3	3188	98.7	619	1 OYR01	guanylate cyclase
4	987	30.5	682	1 OYR02	guanylate cyclase
5	960.5	29.7	685	1 T19968	hypothetical prote
6	955	29.6	686	2 T23721	hypothetical prote
7	938	29.0	751	2 T24458	hypothetical prote
8	914	28.3	699	2 T18984	guanylate cyclase
9	909.5	28.1	732	2 S18325	guanylate cyclase
10	793	24.5	680	1 OYR01	guanylate cyclase
11	770	23.8	691	1 OYB07	guanylate cyclase
12	752.5	23.3	1099	2 T16822	hypothetical prote
13	730.5	22.6	717	2 S23098	guanylate cyclase
14	712	22.0	683	2 JH0810	guanylate cyclase
15	701.5	21.7	583	2 F88642	protein C5A4.3 [1
16	646.5	20.0	947	2 T20156	hypothetical prote
17	481.5	14.9	1061	1 OYH04	natriuretic peptid
18	470	14.5	1057	2 T15720	hypothetical prote
19	468	14.5	1057	1 OYR01	natriuretic peptid
20	466	14.5	1057	2 T15720	natriuretic peptid
21	466	14.4	1057	2 T15720	natriuretic peptid
22	457	14.1	1057	1 OYR01	natriuretic peptid
23	456.5	14.1	632	2 T31667	natriuretic peptid
24	455.5	14.1	632	2 T31666	natriuretic peptid
25	454.5	14.1	1047	1 OYR01	natriuretic peptid
26	453.5	14.0	1047	1 OYR01	natriuretic peptid
27	451	14.0	1525	1 A56699	guanylate cyclase
28	450.5	13.9	1047	2 T45882	guanylate cyclase
29	449.5	13.9	1025	2 T21319	hypothetical prote

30	443	13.7	1056	2 T33167	hypothetical prote
31	438.5	13.6	1110	1 S55279	guanylate cyclase
32	438.5	13.6	1112	2 T28082	hypothetical prote
33	431.5	13.4	1005	2 S33525	guanylate cyclase
34	429	13.3	1012	2 T24384	hypothetical prote
35	428.5	13.3	1018	2 A55915	guanylate cyclase
36	425	13.2	1050	2 S45636	natriuretic peptid
37	424	13.1	1068	2 T42382	guanylate cyclase
38	424	13.1	1108	2 T59385	guanylate cyclase
39	422	13.1	1102	2 JH0717	guanylate cyclase
40	420	13.0	1108	2 B55915	guanylate cyclase
41	419	13.0	1103	2 JC5581	guanylate cyclase
42	418	12.9	1122	2 T28130	hypothetical prote
43	416	12.9	1130	2 A89130	protein F52E1.4 [1
44	415.5	12.9	540	2 T34187	hypothetical prote
45	415.5	12.9	1110	2 I59370	guanylate cyclase

## ALIGNMENTS

### RESULT 1

S23097 guanylate cyclase (EC 4.6.1.2), soluble, 70K chain - human

C/Species: Homo sapiens (man)

C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 20-Apr-2000

C/Accession: S23097

R/Guill, G.; Scholl, U.; Bulle, F.; Gueliaen, G.  
FEBS Lett. 304, 83-88, 1992

A/Title: Molecular cloning of the cDNAs coding for the two subunits of soluble guanyl

A/Reference number: S23097; MID:92316204; PMID:1352257

A/Accession: S23097

A/Molecule type: mRNA

A/Residues: 1-619 <GID>

A/Cross-references: EMBL:X66533; NID:931685; PIDN:CAA7144.1; PID:931686

C/Genetics:

A/Gene: GDB:GUCYB3; GUCYB3; GC-SB3

A/Cross-references: GDB:141992; OMIM:139397

A/Map position: 4q31.3-q33

C/Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

C/Keywords: heterodimer; phosphorus-oxygen lyase

F/373-007/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match	100.0%	Score 3231	DB 1	Length 619
Best Local Similarity	100.0%	Pred. No. 1.1e-215		
Matches 619	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MGFVNHAELELVIRNYGPEWEDIKKEAQLDEBGFVRIIYDSTYDLVAASRYLN 60		
DB	1	MGFVNHAELELVIRNYGPEWEDIKKEAQLDEBGFVRIIYDSTYDLVAASRYLN 60		
QY	61	LNNGEILOMFGKKEFYVCOESGYDTLRVGSNVRELOLDLHDLATYGMRAPSF 120		
DB	61	LNNGEILOMFGKKEFYVCOESGYDTLRVGSNVRELOLDLHDLATYGMRAPSF 120		
QY	121	RCTDAEKGKGLIHYYSERGLDVIIGITTAQAQIHGETIDMKVYIQGNEEDHTQFL 180		
DB	121	RCTDAEKGKGLIHYYSERGLDVIIGITTAQAQIHGETIDMKVYIQGNEEDHTQFL 180		
QY	181	IEEKESKEEDFYEDLDRENGTOESRISPTTCKAPPHIIFDRDLVTVQGNATYRVL 240		
DB	181	IEEKESKEEDFYEDLDRENGTOESRISPTTCKAPPHIIFDRDLVTVQGNATYRVL 240		
QY	241	POLQPGNCSTLVSFLVRPHIDISFHGILSHINTVFLRSKEGLDVEKLECEDELTGTE 300		
DB	241	POLQPGNCSTLVSFLVRPHIDISFHGILSHINTVFLRSKEGLDVEKLECEDELTGTE 300		
QY	301	ISCLRKGQMYIPEADSIIFLCSVYMNDDTLRGLYSDIPLHDATRDVLLGQFR 360		
DB	301	ISCLRKGQMYIPEADSIIFLCSVYMNDDTLRGLYSDIPLHDATRDVLLGQFR 360		
QY	361	EEKTLQDELTLDRDLTLRALEDEKKKDTLTYSLTPSVANELHKKRPVPAKRIDNY 420		
DB	361	EEKTLQDELTLDRDLTLRALEDEKKKDTLTYSLTPSVANELHKKRPVPAKRIDNY 420		

Db 361 EEKLTQLELITDRLOLTLRALDEKKKTDLLYSVLPSPVANELRRKRPVPAKRYDNV 420  
QY 421 TILFSGIVGFNAFCSKHAASGEGAMKIYNLNDLYTRFDLTDSRKNRPVYKVEYVGDYKM 480  
Db 421 TILFSGIVGFNAFCSKHAASGEGAMKIYNLNDLYTRFDLTDSRKNRPVYKVEYVGDYKM 480  
QY 481 TVSGIPEPCIHANASICHALDMMELIAGOVVDGSEVOITIGIHTEGVVTVIGORMPRY 540  
Db 481 TVSGIPEPCIHANASICHALDMMELIAGOVVDGSEVOITIGIHTEGVVTVIGORMPRY 540  
QY 541 CLFGNTVLTSTRTEFTGEGKINSEYTYRCLMSPENDPOPHLEHRGPVSMKGRKPEMO 600  
Db 541 CLFGNTVLTSTRTEFTGEGKINSEYTYRCLMSPENDPOPHLEHRGPVSMKGRKPEMO 600  
QY 601 VWFLSRKNTGTEETKODDD 619  
Db 601 VWFLSRKNTGTEETKODDD 619

## RESULT 2

QY070 guanylate cyclase (EC 4.6.1.2), soluble, beta-1 chain - bovine  
N:Alternate names: guanylate cyclase, soluble, 70K chain  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 18-Jun-1999  
C:Accession: S01653; A38759  
R:Koesling, D.; Herz, J.; Gausepohl, H.; Nitromond, F.; Hlansch, K.D.; Boehm  
FEBS Lett. 239, 29-34, 1988  
A:Title: The primary structure of the 70 kDa subunit of bovine soluble guanylate cyclase  
A:Reference number: S01653; MUID:89031214; PMID:2903071  
A:Accession: S01653  
A:Molecule type: mRNA  
A:Residues: 1-619 <KOE>  
A:Cross-references: EMBL:Y00770; NID:9407; PIDN:CA68739.1; PID:9408  
A:Accession: A38759  
A:Molecule type: protein  
A:Residues: 1-6; 28-32; 41-45; 96-109; 337-343; 396-406; 562-569 <KOE2>  
C:Genetics:  
A:Introns: 281/3; 392/2  
C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology  
C:Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase  
F:373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 99.2%; Score 3206; DB 1; Length 619;  
Best Local Similarity 99.0%; Pred. No. 5.8e-214;  
Matches 613; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MYGFVNALDELIVIRNGPEWEDIKKEAQDDEGQFLVRIITYDDSKTYDLVAASRYLN 60  
Db 1 MYGFVNALDELIVIRNGPEWEDIKKEAQDDEGQFLVRIITYDDSKTYDLVAASRYLN 60  
QY 61 LNAGEIILQMFGRMFVFCQESGYDTILRVLSNVREFLQNDALHDHLATITPGMRAPSF 120  
Db 61 LNAGEIILQMFGRMFVFCQESGYDTILRVLSNVREFLQNDALHDHLATITPGMRAPSF 120  
QY 121 RCTDAEKGKGLILHYSEREGIDIVIGIKTVAQOIHGTEIDMKVIOQRNEECDHTQFL 180  
Db 121 RCTDAEKGKGLILHYSEREGIDIVIGIKTVAQOIHGTEIDMKVIOQRNEECDHTQFL 180  
QY 181 IEEKSEKEDFYEDLDREENGTOESRISPTTCKAPFPHIIFRDVLYVQCGNAIYRVL 240  
Db 181 IEEKSEKEDFYEDLDREENGTOESRISPTTCKAPFPHIIFRDVLYVQCGNAIYRVL 240  
QY 241 POLQPGKCSLSVSLVRPHIDISFHGILSHINIVFVLRSGEGLLDVEKLECEDELGTGE 300  
Db 241 POLQPGKCSLSVSLVRPHIDISFHGILSHINIVFVLRSGEGLLDVEKLECEDELGTGE 300  
QY 301 ISCLRKGOMIYLPDASISILFLCSPSVANNDDLRRLGLYLSIDPLHDATRDVLLGEOFR 360  
Db 301 ISCLRKGOMIYLPDASISILFLCSPSVANNDDLRRLGLYLSIDPLHDATRDVLLGEOFR 360  
QY 361 EEFYLTQLELITDRLOLTLRALDEKKKTDLLYSVLPSPVANELRRKRPVPAKRYDNV 420  
Db 361 EEFYLTQLELITDRLOLTLRALDEKKKTDLLYSVLPSPVANELRRKRPVPAKRYDNV 420

Db 361 EEFYLTQLELITDRLOLTLRALDEKKKTDLLYSVLPSPVANELRRKRPVPAKRYDNV 420  
QY 421 TILFSGIVGFNAFCSKHAASGEGAMKIYNLNDLYTRFDLTDSRKNRPVYKVEYVGDYKM 480  
Db 421 TILFSGIVGFNAFCSKHAASGEGAMKIYNLNDLYTRFDLTDSRKNRPVYKVEYVGDYKM 480  
QY 481 TVSGIPEPCIHANASICHALDMMELIAGOVVDGSEVOITIGIHTEGVVTVIGORMPRY 540  
Db 481 TVSGIPEPCIHANASICHALDMMELIAGOVVDGSEVOITIGIHTEGVVTVIGORMPRY 540  
QY 541 CLFGNTVLTSTRTEFTGEGKINSEYTYRCLMSPENDPOPHLEHRGPVSMKGRKPEMO 600  
Db 541 CLFGNTVLTSTRTEFTGEGKINSEYTYRCLMSPENDPOPHLEHRGPVSMKGRKPEMO 600  
QY 601 VWFLSRKNTGTEETKODDD 619  
Db 601 VWFLSRKNTGTEETKODDD 619

## RESULT 3

QY071 guanylate cyclase (EC 4.6.1.2), soluble, beta-1 chain - rat  
N:Alternate names: guanylate cyclase, soluble, 70K chain  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Jun-1999  
C:Accession: A31871  
R:Nakane, M.; Sahel, S.; Kuno, T.; Ishii, K.; Murad, F.  
Biochem. Biophys. Res. Commun. 157, 1139-1147, 1988  
A:Title: Molecular cloning of a cDNA coding for 70 kilodalton subunit of soluble guan  
A:Reference number: A31871; MUID:89087429; PMID:2905128  
A:Accession: A31871  
A:Molecule type: mRNA  
A:Residues: 1-619 <NMR>  
A:Cross-references: GB:M22562; NID:9204273; PIDN:AAA41204.1; PID:9204274  
C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology  
C:Keywords: cGMP biosynthesis; heterodimer; phosphorus-oxygen lyase  
F:373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 98.7%; Score 3188; DB 1; Length 619;  
Best Local Similarity 98.5%; Pred. No. 1e-212;  
Matches 610; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MYGFVNALDELIVIRNGPEWEDIKKEAQDDEGQFLVRIITYDDSKTYDLVAASRYLN 60  
Db 1 MYGFVNALDELIVIRNGPEWEDIKKEAQDDEGQFLVRIITYDDSKTYDLVAASRYLN 60  
QY 61 LNAGEIILQMFGRMFVFCQESGYDTILRVLSNVREFLQNDALHDHLATITPGMRAPSF 120  
Db 61 LNAGEIILQMFGRMFVFCQESGYDTILRVLSNVREFLQNDALHDHLATITPGMRAPSF 120  
QY 121 RCTDAEKGKGLILHYSEREGIDIVIGIKTVAQOIHGTEIDMKVIOQRNEECDHTQFL 180  
Db 121 RCTDAEKGKGLILHYSEREGIDIVIGIKTVAQOIHGTEIDMKVIOQRNEECDHTQFL 180  
QY 181 IEEKSEKEDFYEDLDREENGTOESRISPTTCKAPFPHIIFRDVLYVQCGNAIYRVL 240  
Db 181 IEEKSEKEDFYEDLDREENGTOESRISPTTCKAPFPHIIFRDVLYVQCGNAIYRVL 240  
QY 241 POLQPGKCSLSVSLVRPHIDISFHGILSHINIVFVLRSGEGLLDVEKLECEDELGTGE 300  
Db 241 POLQPGKCSLSVSLVRPHIDISFHGILSHINIVFVLRSGEGLLDVEKLECEDELGTGE 300  
QY 301 ISCLRKGOMIYLPDASISILFLCSPSVANNDDLRRLGLYLSIDPLHDATRDVLLGEOFR 360  
Db 301 ISCLRKGOMIYLPDASISILFLCSPSVANNDDLRRLGLYLSIDPLHDATRDVLLGEOFR 360  
QY 361 EEFYLTQLELITDRLOLTLRALDEKKKTDLLYSVLPSPVANELRRKRPVPAKRYDNV 420  
Db 361 EEFYLTQLELITDRLOLTLRALDEKKKTDLLYSVLPSPVANELRRKRPVPAKRYDNV 420  
QY 421 TILFSGIVGFNAFCSKHAASGEGAMKIYNLNDLYTRFDLTDSRKNRPVYKVEYVGDYKM 480  
Db 421 TILFSGIVGFNAFCSKHAASGEGAMKIYNLNDLYTRFDLTDSRKNRPVYKVEYVGDYKM 480

OY 481 TVSGIPEPCIHANSICHLADMEIAGOVDSVOITIGIHGEVAVTIGORMPRY 540  
 DB 481 TVSGIPEPCIHANSICHLADMEIAGOVDSVOITIGIHGEVAVTIGORMPRY 540  
 OY 541 CLFGTVLTSRTETGKGINSEYTYRCIMSPENDPOFHLHRGPVSKGKKEPMQ 600  
 DB 541 CLFGTVLTSRTETGKGINSEYTYRCIMSPENDPOFHLHRGPVSKGKKEPMQ 600  
 OY 601 VWFSLRKNTGTEETKODD 619  
 DB 601 VWFSLRKNTGTEETKODD 619

#### RESULT 4

querylate cyclase (EC-4.6.1.2), soluble, beta-2 chain - rat  
 N:Alternates names: guanylate cyclase, soluble, 76k chain  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Jun-1999  
 C:Accession: A36228  
 R:Yuen, P.S.T.; Potter, L.R.; Garbers, D.L.  
 Biochemistry 29, 10872-10878, 1990  
 A:Title: A new form of guanylyl cyclase is preferentially expressed in rat kidney.  
 A:Reference number: A36228; MUID:91105012; PMID:1980215

A:Accession: A36228  
 A:Molecule type: mRNA  
 A:Residues: 1-682 <YUE>  
 A:Cross-references: GB:M57507; GB:J05308; NID:9204279; PIDN:AAA41207.1; PID:9204280  
 A:Experimental source: kidney  
 C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology  
 C:Keywords: cGMP biosynthesis, heterodimer, phosphorus-oxygen lyase  
 F:360-584/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 30.5%; Score 987; DB 1; Length 682;  
 Best Local Similarity 36.9%; Pred. No. 2.4e-60;  
 Matches 225; Conservative 104; Mismatches 195; Indels 86; Gaps 12;

OY 66 ILQMGKMFVFCQSGYDTILRLVGSNREFLQMLDHLHATYRGMRAFSRCDTA 125  
 DB 4 ILKLEGEYFFKCKSGYDRMLRTLGNLTERTEMLDLSHSLALSTQEMNAPSRYVEG 63  
 OY 126 EKGGLIHYSERGLDIIYIGITKYAQQIHGEIDMKYIQQRNEC-----DHIOF 179  
 DB 64 ADG-AMLHYSDRGLCHIVPGITLVAKDFPDVAMSLDM-NEEVERTGKEHYVF 121  
 OY 180 LIEEKSEK-----EDYED----- 194  
 DB 122 LVQKAAHQINGAKRSQSGSDSOADBALOGTLIRMKERYLNIPVCPGKSHSTAVRA 181  
 OY 195 -----LDRFEENGTOESRISPYTECKAPFHIIFEDRLVYTQCGNAIYRVLPOQP 245  
 DB 182 SYLFEGKPLRDFQVYPERLMEVEEFCDAPEFHIIVDEALRVQAGNVIQKYVPGILT 241  
 OY 246 GNCSTLSFSLVRPHIDISFGILSHINTVYLRKSEGLLDVEKLECEDELGTETISCLR 305  
 DB 242 QKFALDEYFSIHPQVTNISSICKFINSOFLKTRKEMPRAR-----KSPQMK 292  
 OY 306 LKGMIITVPEADSLIFLSPSYMNLDDLRLRGILYSDIPLHDATRDIVILGROFEEYKL 365  
 DB 293 LKGQIMMESLRCMTFCSPNRSLOLEESKMLSDIAPHDITDILLNORLAEMEL 352  
 OY 366 TOELETLDRLQTLRALEDEKKTDTLLSYLPSVANELRHRKRPVPAKRYDNTYILFS 425  
 DB 353 SCQLEKKKEELRLSNHLAIEKKKTETILYAMLPEHVANQLEKGRVAAAGEETCTILFS 412  
 OY 426 GIVGNAFCSKHASEGAMKYVNLINDLYTRPDTLDSKKNPYYKVVETVGDKYMTVSGL 485  
 DB 413 DVVETFNIC---ACE-PIQIVNMLNSMTSKFDRITSVD---YKVEITIGAYVYVGV 465  
 OY 486 PEPCHHARSICHLADMEIAGOV--VDGESVQITIGIHGEVAVTIGORMPRYCLF 543  
 DB 466 PVPESHQORVANPALGMRISAKEVWNPVTEGPIQIRGSHRGVLAGVAGDMKPRYCLF 525

OY 544 GNTVNLTSRTETGKGINSEYTYRCIMSPENDPOFHLHRGPVSKGKKEPMQVWF 603  
 DB 526 GDTVATASMEHSHGLPSKVLHLSPTAHRAL-----KNKGELIYRROELIEYKGRK-NTIYF 579  
 OY 604 LSRKNTGTEET 613  
 DB 580 LIQNLINATED 589

#### RESULT 5

hypothetical protein C46E1.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T19968; T23092  
 R:Lennard, N.  
 submitted to the EMBL Data Library, November 1997

A:Reference number: T19968  
 A:Accession: T19968  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-685 <WIL>  
 A:Cross-references: EMBL:AL008867; PIDN:CAA15512.1; GSPDB:GN00028; CESP:C46E1.2  
 A:Experimental source: clone C46E1  
 R:Lennard, N.  
 submitted to the EMBL Data Library, October 1997  
 A:Reference number: T19673  
 A:Accession: T23092  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-685 <WIL>  
 A:Cross-references: EMBL:299942; PIDN:CAB17073.1; GSPDB:GN00028; CESP:C46E1.2  
 A:Experimental source: clone H13M06  
 C:Genetics:  
 A:Gene: CESP:C46E1.2  
 A:Map position: X  
 A:Intons: 11/3; 36/2; 70/1; 133/2; 179/1; 202/3; 258/2; 298/2; 301/3; 325/2; 405/2;  
 C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

Query Match 29.7%; Score 960.5; DB 2; Length 685;  
 Best Local Similarity 34.2%; Pred. No. 1.6e-58;  
 Matches 220; Conservative 131; Mismatches 234; Indels 59; Gaps 17;

OY 2 YGFVNAHLALVIRNYGPEWMDIRKEAOLDEGGLVRIIYDDSKTYDLVAASVNLV 61  
 DB 12 RGFHESITQIMIRYIGELFWFSKVLERAGFEAGKENTINHYISDADTESLVAVSYILAV 71  
 OY 62 NAGEILOMGKMFVFCQSGYDTILRLVGSNREFLQMLDHLHATYRGMRA 117  
 DB 72 TREQWEMYGCLFIQYMETGMDLIRSMPLKGLNDLSLHFEIDHV--VYKANLNG 129  
 OY 118 PSFRCTDAKGGILHYSERGLDIIYIGITKYAQQIHGEIDMKYI--QQRNEC- 174  
 DB 130 PSFRCDNDGP-ITLHYTGPRGLYPIYKGVLRKAKVFKLDVSMITTGTRORSVQMA 188  
 OY 175 -----DHIOFLE---EKESKEDEYEDLDPEENGTOESRISPYTECKAPFHIIFR 225  
 DB 189 TGERIEEHYIFLKLINTQSNDEALGTVQHSNN--TKIRLTHDFISTPPIYAVYDQ 246  
 OY 226 DLVATQCGNAIYRVLPO--LQPGNSLLSVSLVRPHIDISFGILSHINTVLRSEGL 284  
 DB 247 DCKIVQVGRLEYNHPIKDLISVGTPLMRIFEVTRPQIPLDPSICNFINAVFVLYQVTP 306  
 OY 285 LDVER-----LECEDE-----LTGEISCLRKGMIYIPEADSLIFLCSPSV 327  
 DB 307 MEFORANRAAQAIEASENLYEDNNGALALSOCHLKLGQMLMSGCHIMTYLCSPPV 366  
 OY 328 NMIDDLTRGLYSIDPLDADTRDVLJGEOPRFEKTKQGELEITDRLQTLRALEDEK 387  
 DB 367 TSIPELQIGLULIAMPIDHPTDRLILLQQRISDEVEMQLQLEANNQJLENAKDIETVEK 426  
 OY 388 KRTDILLVSLPDSVANELRHRKRPVPAKRYDNTIIFSGIVGFN--AFCSKHASGEGAM 444

Db 427 GKTDLALREMLPPSVAAQLKQGLSVKAEKEEATVMTDPTTQOIVPLCT-----PK 479  
 445 KIYNLNDLYTRFDLTDSRRKNPFVYKVTGDKMYVSGLPKPCIHHSRICHALDM 504  
 480 DIHLNLELFTKDRGLGLOK---AKVVEVGDSSYNKVGIPDLVDHCHVICHALGMV 536  
 505 EINGOV--QVDGESVOITIGIHGEVVTGVIQGMMPRYCLFGMTVNLTSRTTGGKGT 562  
 537 MEKRYVCDPTTNTPLHFRAGIHSGPVGAVGAMPRYCLFGDPTVNTSRMESHPIGRI 596  
 563 NVESEYTRCLMSPENDPOFHLHRGPVSKGKKEPQWVLSR 606  
 597 HSENAKRC-----AESTGRFEFPRGRVQIKGGE-WNTYFLNR 635

RESULT 6  
T23721

hypothetical protein M04G12.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 31-Jan-2000

C/Accession: T23721

R: Sims, M.

submitted to the EMBL Data Library, October 1996

A/Reference number: Z19789

A/Accession: T23721

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-686 <MWL>

A/Cross-references: EMBL:281103; PIDN: CAB03210.1; GSPDB: GN00023; CESP: M04G12.3

A/Experimental source: clone M04G12

C/Genetics:

A/Map position: 5

A/Intons: 1/3; 26/2; 60/1; 192/3; 255/2; 321/2; 401/2; 645/1

C/Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

Query Match 29.68; Score 955; DB 2; Length 686;  
 Best Local Similarity 34.28; Pred. No. 4e-58;

Matches 230; Conservative 126; Mismatches 235; Indels 82; Gaps 20;

QY 1 MYGFVNAHLELVIRNYGPEWEDIRKEAQDDEGQFLVRIITDSTKYDLVAASKRYLN 60  
 1 MGFVHESRQVLTIRKYGEDVYLQVLERSGFENGKEIVNHYISDIDTVLVDSVSLVK 60  
 61 LNAGEILQMGKMFVFCOESGYDTILRVYGSNVEFLQNLDAH--DHLATY-PGMR 116  
 61 VTRDQIWEVYGFLLTYSMELGMDLVRSKSPMLKGLMDLSLHYFIDHV--VYKANLR 118  
 117 APSFCTDAEKGGILHYSEREGLODIYIGIITVAQOIHGTEIDMKVYIQORNEEC- 174  
 119 GPFRCFENPDGT-LMLHYFTGRPGLYHYGVAKVAKLVFMDISL-VVQGRTOHSV 176  
 175 -----DHQFLIEKE-----SKEDFYEDLDREENGTOESRISPY 211  
 177 MNGEVEEVEHVIILINVEERRRDSSTISALTSPVDPGEIID-----DNKAKSLQ 229  
 212 TFCKAPPHIIFRDLYVTCGNATYRVLPQ--LQPGNCSLSVSFLVRPHIDISFGL 269  
 230 DFRALRYHVEFLDESCRLLVQCGDELYNHLIPNELLOGT-PILRIFEINRQIDLPENIC 288  
 270 SHINTVFVLSKRGGL-----DVKEKCEDELTGTEISL-LRKGMQIVL 313  
 289 NFINAVFVLOVKTSPLRKHKHMANMTEKEQVEAEVEEVESENELTQGHKLKGMQML 348  
 314 PEADSLIFLSPVNMNDLTRGLYSLDIPLDATRDVLLGQFEEKTLQOELEILT 373  
 349 STKHIIYICSPVYTGINELMORGMLTANPLADIRDLILLQORISDVENVLQLEANN 408  
 374 DRLQTLRALDEKRRKTDLLYSVLPSPVANELRHRKRPYAKRYDNTVILFSGIVGNAP 433  
 409 EOLETWTHELEVROKTDLSILKMDLPRIKAKOLLGSHLEPCYE-ATVAFCOLPAPFOI 467

QY 434 ---CSKHASEGAMKIVNLNDLYTRFDLTDSRRKNPFVYKVTGDKMYVSGLPKPC 490  
 468 IPVCQ-----PRNIKLLNVEFKLDIRIVLRG--YKVEVSDSYVSGIDPYTS 517  
 491 HHARSICHLADMMELTAGOVQ--VDGESVOITIGIHGEVVTGVIQGMMPRYCLFGMTV 548  
 518 EHNENMCHVALGMMMEKRSYMDPNKTPILLRGLHSGTIIAGVGVGKMPRYCLFGMTV 577  
 549 LTRTETGGKGINSEYTRCLMSPENDPOFHLHRGPVSKGKKEPQWVLSR-- 606  
 578 LASOMESLGVAGNIQCSSWY-----SKAMTGFEEPSPRGRINVRKRGD-VETIFLMRSL 632  
 607 KMTGETEKODD 619  
 633 KKSIMEITDHERD 645

RESULT 7  
T24458

hypothetical protein T04D3.4 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000

C/Accession: T24458

R: Kershaw, J.

submitted to the EMBL Data Library, October 1996

A/Reference number: Z19893

A/Accession: T24458

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-751 <MWL>

A/Cross-references: EMBL:281114; PIDN: CAB03288.1; GSPDB: GN00019; CESP: T04D3.4

A/Experimental source: clone T04D3

C/Genetics:

A/Map position: 4

A/Intons: 1/3; 26/2; 95/3; 149/1; 213/3; 247/3; 291/3; 332/3; 369/2; 418/1; 462/3; 4

C/Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

Query Match 29.08; Score 938; DB 2; Length 751;  
 Best Local Similarity 35.08; Pred. No. 6.8e-57;

Matches 220; Conservative 125; Mismatches 236; Indels 48; Gaps 15;

QY 1 MYGFVNAHLELVIRNYGPEWEDIRKEAQDDEGQFLVRIITDSTKYDLVAASKRYLN 60  
 1 MGFVHESRQVLTIRKYGEDVYLQVLERSGFENGKEIVNHYISDIDTVLVDSVSLVK 60  
 61 LNAGEILQMGKMFVFCOESGYDTILRVYGSNVEFLQNLDAH--DHLATY-PGMR 116  
 61 IPLEIWEAVGFLIQFTMETGMDLIRAMAPDLGFLSDLSLHYFIDHV--VYTKLR 118  
 117 APSFCTDAEKGGILHYSEREGLODIYIGIITVAQOIHGTEIDMKVYIQORNEEC- 175  
 119 GPFRCFENPDGT-LMLHYFTGRPGLYHYGVAKVAKLVFMDISL-VVQGRTOHSV 176  
 176 -----DHQFLIEKE-----SKEDFYEDLDREENGTOESRISPY 211  
 178 FTEHVAFFVLTQLEMANSTQPKSISSKADQIDL---STGIYE--ISSDSFLAPYH 231  
 222 IFDRDLVYTCGNATYRVLPQ--LQPGNCSLSVSFLVRPHIDISFGL 281  
 232 CFDPDLFVHFHFGNFIKTPNARMQETRVLDLELVHPEVPFESIKYKNSLAFPRK 291  
 282 EGLDVEKCEDELTGTEISCLRLKGQMTYLPKADSLIFLSPVNMNDLTRGLYSL 341  
 292 -GLGDIY-----HNANDEKTYLLKGSWFIIDEGRYIILYMCVNTVYVLELIERMHL 344  
 342 DIHLNATRDVLLGQFEEKTLQOELEILT 373  
 345 DMQRHGTBDVIMLNOSRMSQVELNRTLEETTKKLKMAQOELEIKOKDELCELMPS 404  
 402 VANELRHRKRPVAKRYDNTVILFSGIVGNAP 433  
 405 VADSLSGRAMDAKFAADCTILFTDITVFTNICAMCTPYD---VTLNDLILNRDRLV 460

QY 462 DSKNPFYKTYGVGDKMTYVSGLEPDCIHARRSICHLALDMEIAGOV--OVDESVOI 519  
 DB 461 GLHD---AYKVEITGDAMVIGVPERCENHAEPLNTISGIMLESKIVLSPITFKPIKI 517  
 QY 520 TIGIHTEGVYGVIGORHPRCLFGNTVNLSTRTTEGKINSEYTYRCLMSPEND 579  
 DB 518 RLGHGCPVAVGVGKIKPRCLFGDYVNAVAKMESNGIOCKIHSE--TGKILGLAN 574  
 QY 580 PQFHLHRRGPVSMKGRKEPMQVWFLSRKN 608  
 DB 575 PSYVFIDRGNTEIRG-KGMATYFLERND 602

## RESULT 8

118984  
 hypochelical protein C06B3.8 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C:Accession: T18984  
 R:Pericy, C.  
 submitted to the EMBL Data Library, July 1996  
 A:Reference number: Z19056  
 A:Accession: T18984  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-699 <M1>  
 A:Cross-references: EMBL:Z77652; PIDN:CA801118.1; GSPDB:GN00023; CESP:C06B3.8  
 A:Experimental source: clone C06B3  
 C:Genetics:  
 A:Gene: CESP:C06B3.8  
 A:Map position: 5  
 A:Introns: 1/3; 60/1; 192/3; 210/2; 270/2; 336/2; 415/2; 659/1  
 C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

Query Match 28.3%; Score 914; DB 2: Length 699;  
 Best Local Similarity 32.9%; Pred. No. 2.8e-55;

Matches 219; Conservative 132; Mismatches 234; Indels 80; Gaps 19;

QY 1 MYGFVNAHLELVIRNRYPEWEDIKKEAQDDEGQVIRIYDSKTYDLVAASKVYN 60  
 DB 1 MFGFTHESIRQIVINRYGEDMTQYVLESGFESGKENTMNYSDTDYVLDVSIVLK 60  
 QY 61 LNAGEIIMFGKMFVFCQESGYDTILRYLGSNVRFLQNLDAH--DHATYV-PGMR 116  
 DB 61 VTKQVMEYMGGLFTTYSMETIGMDLVSRMSPNLKFIDNLSLHYFIDH--VYKANLR 118  
 QY 117 APSFCTDAEKKGILHYISREGLQDIYIGIITVAQIHGTEIDMKVIOQRNEC-- 174  
 DB 119 GPFCEETPDGT-LTLHYFTGRPLHYHVKGVKAVKARVPDLDTL--VVGKTRQSVYH 176  
 QY 175 -----DHQFLIEKESKEDEYEDL-----DREENGTOES----- 206  
 DB 177 MNGSRVEHVEFLINKKGGQFHPFLAIIVYGESEPRRSESEVSLSTNANFTIYD 236  
 QY 207 ---RISPTTCKAPFPHIIFRDLYVTCGNATRYLPQ--IQPGNCILSVFSVLRPHI 261  
 DB 237 DTGLSLDQFSKALPHEVYHEDISCKLYOCSELSHNPMLLPQGT-PILRFFEINRPOI 295  
 QY 262 DISFGLSHINVTYVLRKSGELD-----VEKLECEDELSTETIS---CUL 306  
 DB 296 PLDFENICNFINAVVLOVKTSPLLKKHMDAMSQELKQEMETLDEDAITNELAQHHL 355  
 QY 307 KGOMIYLRPADSILFLCSPSVNMDLTLRRGLYSLDIPLDATRDLYVLGPQFREYKLT 366  
 DB 366 KGOMMLASKKHIIIVCSPTYVSINELMQGRITAMPPLHATRDILINQRLSDVEYN 415  
 QY 367 QELELTIRLOLTALADEKKKTDTLYSVLPSPYANELRKRPPAKRYNVTILFSG 426  
 DB 416 IQLENNPDLTMTRELEERQKTDISLMDLPRLIAQQLSGEHLAECEHE-ATVMPCD 474  
 QY 427 IYGFNAF---CSKHASGEGAMKIYVNLNDLYTRFDTLSRKPNFYKTYGVGDKMTVS 483

DB 475 LPAFOALPQCS-----PKDIYMLNEIFKKLDRIYVIRG---VYKVEYVSDSYMAVS 524  
 QY 484 GLPEPCIHARRSICHLALDMEIAGOV--OVDESVOITIGIHTEGVYGVIGORHPRYC 541  
 DB 525 GIDPOTPREAHENMCMHVALGMMMEARSVIDPVSKTFFLRIGHSGITAGVGYVHPKXC 584  
 QY 542 LFGNTVNLSTRTTEGKINSEYTYRCLMSPENDPQFHLHRRGPVSMKGRKEPMQV 601  
 DB 585 LFGETVTLASQMESLGMAGKIOCSKWAYOKAM-----ETGRFEFSPGRIDVR-QRGITET 639  
 QY 602 WFLSR 606  
 DB 640 YFLTR 644

## RESULT 9

518325  
 guanylate cyclase, soluble, alpha chain - human  
 C:Species: Homo sapiens (man)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Jul-1999  
 C:Accession: S18325  
 R:Harteneck, C.; Medel, B.; Koesling, D.; Malkewitz, J.; Boehme, E.; Schultz, G.  
 FEBS Lett. 292, 217-222, 1991  
 A:Title: Molecular cloning and expression of a new alpha-subunit of soluble guanylyl  
 A:Reference number: S18325; M01D:92070494; PMID:1683630  
 A:Accession: S18325  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-732 <HRA>  
 A:Cross-references: EMBL:X63282; NID:g31670; PIDN:CA444921.1; PID:g31671  
 C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology  
 F:473-701/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 28.1%; Score 909.5; DB 2: Length 732;  
 Best Local Similarity 37.7%; Pred. No. 6.1e-55;

Matches 208; Conservative 111; Mismatches 175; Indels 57; Gaps 14;

QY 51 LVAAASKVNLNAGEIIMFGKMFVFCQESGYDTILRYLGSNVRFLQNLDAHDLAT 110  
 DB 154 ILQCVANIILGLFEEIQRKRGEEFNICFEN-ERYLRAVGGILODFNGFALLEHINT 212  
 QY 111 IY---PGMRAPSFCTDAEKKGILHYISREGLQDIYIGIITVAQIHGTEIDMKVY 167  
 DB 213 SFGRQATLESPSFLKEPEGT-LMLHYHPHHIVGFAMGIMKAGKITY--RLDVEYE 269  
 QY 168 QQRNE-----CDHTQFLIEKESKEDEYEDLDLDRFENGQESRISPYRCKAF 217  
 DB 270 QVANERKLCSDVSNPGNSCLTFLI--KECENTNIMKNLPQGTQVPADRLISINTFCRAF 327  
 QY 218 PHIIIFRDLYVTCGNATRYLPQLPNGNSILSV-----FSLVRPHIDISFHGLISH 271  
 DB 328 PHLMDPMSVYLQEGELRKOL-----RCQTHKVLKREDFEIVSPVYNAFTFEVLLR 381  
 QY 272 INTVYVLRKSGELDVYEKLECEDELTEGTE--ISCLRLKQMIYLRPADSILFLCSPSVAN 329  
 DB 382 LSTPEVIRK-----PEAGSSKNKQVMKQMHVPESSNIIILFGLSPCYVK 429  
 QY 330 LIDLTRGLYSLDIPLDATRDLYVLGEPFREYKTOLELTITDLQTLALADEKKK 389  
 DB 430 LDELMGRGLHLDIPRHATRDYVLYGEGAKQDGLKRRMDIKATLETHQALEKKK 489  
 QY 390 TDTLYSVLPSPYANELRKRPPAKRYNVTILFSGIVGFNAFCSKHASGEGAMKIYVNL 449  
 DB 490 TYDLYAISIPGVAOQLMOGOOQVAKRFDVTMLFSDIYGFALICAOQ---PMQVISM 545  
 QY 450 LNDLYTRFDTLSRKNPF--YKRYTVDKMTVSGLPEPCIHARRSICHLALDMEIAG 507  
 DB 546 LNELTYRFD-----HOCGLDITKVTETIDATCAVAGLRKRSICAKKPTALNALKMEIS 600  
 QY 508 GGOV--DGESVOITIGIHTEGVYGVIGORHPRYCFLGNTVNLSTRTTEGKINSEY 566  
 DB 601 EETLPDGRPIQRIISIHSGSVLAGVYGRMPRYCLFGNNVTLASAFESGSHRRIRINVP 660

## RESULT 10

OY 567 YTYRCIMSPEN 577  
DB 661 TTYQLKREBS 671

OYRA1

guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 chain - rat  
N:Alternate names: guanylate cyclase, soluble, 77k chain  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 18-Jun-1999  
C:Accession: A38297  
R:Nakane, M.; Arai, K.; Saheki, S.; Kuno, T.; Buechler, W.; Mured, F.  
J. Biol. Chem. 265, 16841-16845, 1990  
A:Title: Molecular cloning and expression of cDNAs coding for soluble guanylate cyclase  
A:Accession: A38297; MUID:91009100; PMID:1698769  
A:Molecule type: mRNA  
A:Residues: 1-690 <NAK>  
A:Cross-references: GB:M57405; GB:M36075; NID:g204277; PIDN:AAA41206.1; PID:g204278  
C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology  
C:Keywords: GMP biosynthesis; heterodimer; phosphorus-oxygen lyase  
F:432-660/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 24.58; Score 793; DB 1; Length 690;  
Best Local Similarity 34.78; Pred. No. 6.5e-47;

Matches 214; Conservative 100; Mismatches 215; Indels 88; Gaps 20;

OY 27 KEAQLDEGQFLVRIYDSDKTYDLVAASKVLNLNGEILQ-MFGKMFVFCOESGYDR 85  
DB 102 KERNNSEKEDLERIAEA-----IAAGVPV-----EVLKDSLGELEFKICEED-EH 149  
OY 86 ILNVLGSNVEFLQNDAL-----HDHLATVPGMARPSFCTDAEKGLILHYSERE 140  
DB 150 ILGVGGTLDLDFNSFTLLKSSHCOEAERGRLEDAISILCDKQD-DELYNYFFPKR 208  
OY 141 GLGDVIGIKTVAAOIHGTEDMKVYIQ-RNECDHTOFL-----LEEKSEKEDF 191  
DB 209 TVALLPGLITKAARILYSEHVEVSLMPCFRSEC--TEFVNOPYLLSYHVSSTKPS-- 264  
OY 192 YEDLDREENGTOESRISPYT-ECKAPFHIIFDRDLVTVQGNALYVLPOLQ-PNCS 249  
DB 265 -----LSPGKQSSSLVIFPSLFCCTPFPHMLDRDLAIQLGNGIRLVNKRDFQKPN 318  
OY 250 LLSVPSLVPRHIDISFGILSHINIVYVLSKEGLDVEKLECEDELGTETISCLRIKQ 309  
DB 319 FEEFFELTPRIKIQTFSGIMTMNQFVIVRRW-----DNLVYKSSRYMDLKGQ 368  
OY 310 MIVLPKADSLIFLCSFSVYMLDLTRGLYISDIPLDATRDLYLGEORFEYKILNQEL 369  
DB 369 MIYIVESSALIFLFGSCVDRLDEFTGRLGLISDIPINLADVYLIGEQAARADGLKRL 428  
OY 370 EILTDRLQILRALDEDEKKTDTLLYSVLPSPVANELRHRPVAKRYDVWTLIFSGIV 429  
DB 429 GKILKATLEHQALEEKKTVDLCSIFPSEVAQAQMOQIYQAKKFENVTLFSDIVG 488  
OY 430 FNAFCSKSHASGEGAKMVLNLDLYTRFDLTLDSKKNPFYKYVETGDKMTVSGLEPC 489  
DB 489 FTALICQSCS---PLQVITMLNALYTRFDQCGELD---YYKVELTIGDACAIVAGLRES 541  
OY 490 IHARSICHLALMELIAGOV-QVDESVQITIGIHGEVYVTVIGORMRCLFEGTVN 548  
DB 542 DTHAVQIALMALMELISNEVSPHGPIMRIGLSGVSFAVAGVAKMRYCLFGNNVT 601  
OY 549 LVSRTETTGKRIKINSEYVYRCL-----MSPENSDFQF-----HL-----EH 586  
DB 602 LANKESSCVPRKINVSPTTYRLKDCPGVFTPRSEELPRNFPSPDIPGICHTLAYQH 661  
OY 587 KGPVSMKGKKEPMQVMP 603  
DB 662 QGPNSS-----KPMF 670

## RESULT 11

OYB077

guanylate cyclase (EC 4.6.1.2), soluble, alpha-1 chain - bovine  
N:Alternate names: guanylate cyclase, soluble, 77k chain  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 18-Jun-1999  
C:Accession: S10713; A38767  
R:Koesling, D.; Harteneck, C.; Humbert, P.; Bosserhoff, A.; Frank, R.; Schultz, G.; B  
FEBS Lett. 266, 128-132, 1990  
A:Title: The primary structure of the larger subunit of soluble guanylyl cyclase from  
A:Reference number: S10713; MUID:90306336; PMID:1973124  
A:Accession: S10713  
A:Molecule type: mRNA  
A:Residues: 1-691 <KOE1>  
A:Cross-references: EMBL:X54014; NID:g405; PIDN:CAA37960.1; PID:g406  
A:Accession: A38767  
A:Molecule type: protein  
A:Residues: 118-133; 226-232; 286-293; 319-330; 412-417; 557-571; 629-637 <KOE2>  
C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology  
C:Keywords: GMP biosynthesis; heterodimer; phosphorus-oxygen lyase  
F:434-662/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 23.88; Score 770; DB 1; Length 691;  
Best Local Similarity 33.98; Pred. No. 2.5e-45;

Matches 205; Conservative 90; Mismatches 200; Indels 110; Gaps 16;

OY 27 KEAQLDEGQFLVRIYDSDKTYDLVAASKVLNLNGEIL-QMFGKMFVFCOESGYDR 85  
DB 104 KERNKSLERDFKRIYVQA-----IAAGVPV-----EIKKSLGELEFKICEED-EV 151  
OY 86 ILNVLGSNVEFLQNDALHDHLATVPGMARPSFCTDAEKRG- 130  
DB 152 ILGVGGTLDLDFNSFTLLKSSHCOEAER-KGREDAISILCDKDPD 199  
OY 131 -LILHYSEEGKODIYIGIKTVAAOIHGTEDMKVYIQORNECDHTOFLIEKSKKEE 189  
DB 200 VLIVYVFFPKRITSLIPGIIKKAARILYSEHVEVSTPR----- 240  
OY 190 DFEYEDLDREENGTO-----ESRISPYT-ECKAPFHIIFDRDLV 229  
DB 241 -FHQDCREFVDQCELYSVHRSARPHPPGKPVSSLVIPASLFCCTPFPHMLDRDMSI 299  
OY 230 TQCGNMIYRVLPLQ-PGNCSLSVPSLVPRHIDISFGILSHINIVYVLSKEGLDVE 288  
DB 300 LQLEHGIRRLMSRDYQKPHFDEYFETLTPKISQTFSGIMTMNQFVIVRRW----- 354  
OY 289 KLECEDELGTETISCLRLKQMTYLPKADSLIFLCSFSVYMLDLTRGLYISDIPLDHA 348  
DB 355 -----DMSMKSSRYVDLQKQMTYVMESSILFLGSPCVDRLEDFTGRLGLYISDIPHN 409  
OY 349 TRLVVLIGEORFEYKILQLELITDRLQILRALDEDEKKTDTLLYSVLPSPVANELRH 408  
DB 410 LRLVVLIGEQAARADQDKRLKRLKATLEQAHLAEERKRYVDLCSIFPSEVAQRLW 469  
OY 409 KRPVPAKRYNMTILFSGIVGFNAFCSKSHASGAKIYVNLDTTRPDITLDSKKNP 468  
DB 470 GHAVQAARFEGVNTMLSDIVGFALICQSCS---PLQVITMLNALYTRFDQCGELD--- 522  
OY 469 VYKVEYGGDYMTVYVSGLEPCIIHARSICHLALMELIAGOV-QVDESVQITIGIHTE 527  
DB 523 VYKVEYIGDGYCYAAGLHKRSDTHAVQIALMALMELISHEVYSPGEPITKMGILHS 562  
OY 528 VYTGVIQORMPRYCLFGNTVNLNLSRTETGKRIKINSEYVYRCL-----MSPEN 577  
DB 583 VFAGVGVGKMPRYCLFGNNVTLANKESSCVPRKINVSPTTYRLKDCPGVFTPRSREE 642  
OY 578 SDPQF 582  
DB 643 LPPNF 647

## RESULT 12

T16822

hypothetical protein T07D1.1 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*

C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C/Accession: T16822

R:leimac, D.

submitted to the EMBL Data Library, November 1995

A/Description: The sequence of C. elegans cosmid T07D1.

A/Reference number: Z18584

A/Accession: T16822

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1099 <LEI>

A/Cross-references: EMBL:U41531; NID:g1109802; PID:g1109803; PIDN:AAA83158.1; CESP:T07D1

C/Genetics:

A/Gene: CESP:T07D1.1

A/Introns: 26/3; 94/3; 144/2; 200/3; 239/2; 278/3; 344/1; 382/2; 425/3; 474/3; 523/3; 61

Query Match 23.3%; Score 752.5; DB 2; Length 1099;

Best Local Similarity 29.9%; Pred. No. 8e-44; Mismatches 254; Indels 117; Gaps 24;

Matches 211; Conservative 124; Mismatches 254; Indels 117; Gaps 24;

1 MYGFVNAHLELVIRNYPVEWEDIKKEAOLDEGQFLVRIYDSTKYDVAASKYLN 60

1 MYGLIDHIAIYIKKYESEWSEVKFVSGYTD-TFQMDKKFSGLSHKLIMACHDTG 59

61 LNAGEIILMGKMFVFCQESGYDT-----TLNVLGNSVREPLQNDLADHDLAT 110

60 DPVDELMTNIGTSFKEFLTFEKNVYVLFYIPYLVNLGRTFQPLNGLDNLHEYLRF 119

111 IYPGRARSPFCTDAEKKGILHYSSREGIADIVIGIKYVAQOIHGTEDMAYIQ-Q 169

120 TFPKLPSPFYC-EHESRTGLMTRSKRGFLHYVQOIRNISELQTEVAILDIE 178

170 RNECDHTQFLIEEKSEKEDFYEDLDREENG-T-----QESRISPTFKAPFE 219

179 HDLMEHYIMRLH-----FNNLD-FNNGSTAVRLNDSILEKVKITSDIFDIPE 228

220 HIIDRDILVYQCGNAITRYLPOLQPGNCSLSVSLVPRH-----DISP-- 265

229 IIVRGRKIRNITIGILRVNAGIV-GKKNQDTFLMRPFIRFMEVEYRGTRNRYSPDI 286

266 -----HGILSHINTVFLNS-----KRGILDVKELEDELT 297

287 ACDCFRNFAPSKKRVKSVKSLMLHNNIFELISSDPIODEEGIL-VYTTVDVDDT 345

298 -----GTETICRLKGMITYLEADSLFLCSPSVANNLDDLTRGILYSLDIPLDAT 349

346 EERHMGDEGRKFLSLGOMFYMESICFVGIPVSHLPQMKSGILFINDFALHDS 405

350 RDVILGQFREFRYLTQBELIDNR-LQITRLALEDEKKTDTLLSYLPPSVANELRH 408

406 RDVYLASTQSAELKLLHQEAQSRNRMENNRKERRRTDKLLYQMLPKPSVANOQLH 465

409 KRPPVA--KRYDNVTILFSGIVGFNAFCSKHAASGAKIYVLLNDLYTRFPLDSRN 466

466 GESAVACCERDSVITLFTDIEFPMKSSLT-----PLEVIEFLAVITINFDKIIDTNG- 520

467 PFYKVFETVGDKMYTSGLPEPCIHARSICHLADMEIAQ-VQVDE--SVQITIGI 523

521 --VYKVEITGDAYVVSAGPTETEDAEFLIDCASQIYVEAGKVMNMNKKIKHIDIRAGV 578

524 HTGEVYTVIGQRMRYCLFGNTVNLSTRTETGKGINVSEYTRYCLMSRNDP-QF 582

579 HSGSVAVAGVGLSMRYCLFGTYVANKMEONSSPMILVSETHNKK--EESDPGLY 635

583 HLEHRGPVSMKKKEPMQVFL-----SRKNTGTEERKODD 618

636 QFERRELEIK-DDQITOTFFVVSRRGRVPSFPMNC--ESRQDD 677

C/Species: *Homo sapiens* (man)

C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Jul-1999

C/Accession: S23098

R:Gull, G.; Scholl, U.; Bulle, F.; Guellaen, G.

FEBS Lett. 304, 83-88, 1992

A/Title: Molecular cloning of the cDNAs coding for the two subunits of soluble guanylyl

A/Reference number: S23097; MUID:92316204; PMID:1352257

A/Accession: S23098

A/Molecule type: mRNA

A/Residues: 1-717 <GID>

A/Cross-references: EMBL:X66534; NID:g31663; PIDN:CAA47145.1; PID:g31684

C/Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology

C/Keywords: heterodimer; phosphorus-oxygen lyase

F/430-658/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match 22.6%; Score 730.5; DB 2; Length 717;

Best Local Similarity 36.9%; Pred. No. 1.5e-42; Mismatches 182; Conservative 80; Mismatches 178; Indels 53; Gaps 10;

119 SFRCTDAKKGKGLILHYSSREGIADIVIGIKYVAQOIHGTEDMAYIQQRNECDHTQ 178

187 SILCKDE-DDELHYTFEPKRTSLILPGIIKAAHYLYEVEVSLM---PCGFIND 241

179 -----FLIEKSKEDFYEDLDREENGTOSSRISPYT-FCKAPFRIIDRD 226

242 CSEFNOPYLAYSVMKSTKPS-----LSPKQSGLYLPTSLFCKTFEFHMFMD 293

227 LVVTCQNAIYVLPOLQPGNCSLSVSLVPRHIDISPHGILSHINTVYRSKGLLD 286

294 MIIIDFGNGIRLMNRDQGFNREYFELIPKINOTFSGIMTMNMOFVVRWDS 353

287 VEKLECEDELGTETICRLKGMITYLEADSLFLCSPSVANNLDDLTRGILYSLDIPH 346

354 VKR-----SSRVMDLKGMITYESSALIFLGSCVDRLEDGFRGILYSLDIPH 403

347 DATRDVILGQFREFRYLTQBELITRLQITRLALEDEKKTDTLLSYLPPSVANEL 406

404 NALRDVILGQFREFRYLTQBELITRLQITRLALEDEKKTDTLLSYLPPSVANEL 463

407 RKRPRVAPKRYNVTILFSGIVGFNAFCSKHAASGAKIYVLLNDLYTRFPLDSRN 466

466 WQGVVQAKKFSNVMTLSIDYGFALCSQCS-----PLQVITMLNALTFRQCGEID- 518

467 PFYKVFETVGDKMYTSGLPEPCIHARSICHLADMEIAQ-VQVDESVQITIGIT 525

519 --VYKVEITAMPVILGILHKSPTHAVQIALMLKMEISDEVMSPHGEPIKMLGLHS 576

526 GEVYTVIGQRMRYCLFGNTVNLSTRTETGKGINVSEYTRYCL-----MSP 575

577 GSVFAGVGVKMPRCLEGNVTLANKESCVPRKINVSPTTYRLKDCPGVETPRSR 636

576 ENSDPOFHLHRG 588

637 EELPPNFSEIRG 649

#### RESULT 14

JH0810

guanylate cyclase (EC 4.6.1.2) - fruit fly (*Drosophila melanogaster*)

C/Species: *Drosophila melanogaster*

C/Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 26-Aug-1999

C/Accession: JH0810

R:Yoshikawa, S.; Miyamoto, I.; Aruga, J.; Furutachi, T.; Okano, H.; Mikoshiba, K.

J. Neurochem. 60, 1570-1573, 1993

A/Title: Isolation of a *Drosophila* gene encoding a head-specific guanylyl cyclase.

A/Reference number: JH0810; MUID:93203896; PMID:809578

A/Accession: JH0810

A/Molecule type: mRNA

A/Residues: 1-683 <YOS>

A/Cross-references: GB:S57126; NID:g298674; PID:g298675

A/Experimental source: head

A/Gene: dgc 1

RESULT 13

S23098

guanylate cyclase (EC 4.6.1.2), soluble, 81k chain - human









DR EMBL: Y15723; CAA75738.1; -  
 DR PIR: S23098; S23098.  
 DR HSP: P19687; 1AMN.  
 DR Genew: HGNC:4685; GUCY1A3.  
 DR MIM: 139396; -  
 DR Interpro: IPR001054; G\_cyclase.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART: SM00044; CYCC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS0125; GUANYLATE\_CYCLASES\_2; 1.  
 KM Lyase; GMP synthesis; Multigene family.  
 FT DOMAIN 481 608 GUANYLATE CYCLASE.  
 FT CONFLICT 124 127 AAC -> QQS (IN REF. 1).  
 FT CONFLICT 131 184 VIKESIGEVEFKICEEDENILGVGGTLDFINSESTLLK  
 OSSHOCGRGR -> LSKNLVRFKLVTRKMTSLGM  
 LEAPLKFQLOVPEPTEQPLPRKKGQ (IN REF.  
 1).  
 FT CONFLICT 322 322 MISSING (IN REF. 1).  
 FT CONFLICT 529 535 GDACTVA -> AMPTVL (IN REF. 1).  
 FT CONFLICT 678 690 GNAFLKASGID -> ASQFROSIRNLATYPIYKSLG  
 FSLKMCRASESTIGIDG (IN REF. 1).  
 SQ SEQUENCE 690 AA; 77452 MW; DAIEI4ASEI145ICF CRC64;  
 Query Match 100.0%; Score 3593; DB 1; Length 690;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-243;  
 Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFCSTKLKDKITGECFSLAPGQVNESEEAAGSSSEKATVPICODIPEKNIOESLP 60  
 1 MFCSTKLKDKITGECFSLAPGQVNESEEAAGSSSEKATVPICODIPEKNIOESLP 60  
 DB 1 MFCSTKLKDKITGECFSLAPGQVNESEEAAGSSSEKATVPICODIPEKNIOESLP 60  
 QY 61 QRTSRSRYLHTLAESICKLIPPEERLNVALLQRTLAKHKIKESKSLEREDFEKTIAE 120  
 61 QRTSRSRYLHTLAESICKLIPPEERLNVALLQRTLAKHKIKESKSLEREDFEKTIAE 120  
 DB 61 QRTSRSRYLHTLAESICKLIPPEERLNVALLQRTLAKHKIKESKSLEREDFEKTIAE 120  
 QY 121 QAAVAGPVAVIKESIGEVEFKICEEDENILGVGGTLDFINSESTLLKOSHOCEAG 180  
 121 QAAVAGPVAVIKESIGEVEFKICEEDENILGVGGTLDFINSESTLLKOSHOCEAG 180  
 DB 121 QAAVAGPVAVIKESIGEVEFKICEEDENILGVGGTLDFINSESTLLKOSHOCEAG 180  
 QY 181 KRGRLDASITLCKDEDFLHVVYFFPKRTSLIPGIIKAAHVLVETEVEVSLMPPCF 240  
 181 KRGRLDASITLCKDEDFLHVVYFFPKRTSLIPGIIKAAHVLVETEVEVSLMPPCF 240  
 DB 181 KRGRLDASITLCKDEDFLHVVYFFPKRTSLIPGIIKAAHVLVETEVEVSLMPPCF 240  
 QY 241 HNCSEVNPYLLSYHMAKSTPSLSKPOSSLVPTSLFCKTFPPHMFDMKMTILQ 300  
 241 HNCSEVNPYLLSYHMAKSTPSLSKPOSSLVPTSLFCKTFPPHMFDMKMTILQ 300  
 DB 241 HNCSEVNPYLLSYHMAKSTPSLSKPOSSLVPTSLFCKTFPPHMFDMKMTILQ 300  
 QY 301 FNGSIRLMMRRDPFGCKPNEEFEEILTPKINOTFGSITMLNMQFVVRVRMDSYKKS 360  
 301 FNGSIRLMMRRDPFGCKPNEEFEEILTPKINOTFGSITMLNMQFVVRVRMDSYKKS 360  
 DB 301 FNGSIRLMMRRDPFGCKPNEEFEEILTPKINOTFGSITMLNMQFVVRVRMDSYKKS 360  
 QY 361 SRVADLKGMITYVESSAILFLGSPCVRLDEFTGRGLYLSIDIPINHALRDVYLIGBOAR 420  
 361 SRVADLKGMITYVESSAILFLGSPCVRLDEFTGRGLYLSIDIPINHALRDVYLIGBOAR 420  
 DB 361 SRVADLKGMITYVESSAILFLGSPCVRLDEFTGRGLYLSIDIPINHALRDVYLIGBOAR 420  
 QY 421 AOGGLKRRGLKATLEQAHAQALEEKKTVDDLCSIFPEVAQAOLMQGOVQAOKKSNV 480  
 421 AOGGLKRRGLKATLEQAHAQALEEKKTVDDLCSIFPEVAQAOLMQGOVQAOKKSNV 480  
 DB 421 AOGGLKRRGLKATLEQAHAQALEEKKTVDDLCSIFPEVAQAOLMQGOVQAOKKSNV 480  
 QY 481 TMLFSDIVGFTALICSCSPLOVITMLNALTREFDOCGEIDVYKVERTIGDAYCAGSLHK 540  
 481 TMLFSDIVGFTALICSCSPLOVITMLNALTREFDOCGEIDVYKVERTIGDAYCAGSLHK 540  
 DB 481 TMLFSDIVGFTALICSCSPLOVITMLNALTREFDOCGEIDVYKVERTIGDAYCAGSLHK 540  
 QY 541 ESDTHAVQIALMLKMMELSDVMSPHGEDIKMRIGLHSGSVTAVGVVGMPCRYCLEGN 600  
 541 ESDTHAVQIALMLKMMELSDVMSPHGEDIKMRIGLHSGSVTAVGVVGMPCRYCLEGN 600  
 DB 541 ESDTHAVQIALMLKMMELSDVMSPHGEDIKMRIGLHSGSVTAVGVVGMPCRYCLEGN 600  
 QY 601 VTLANKFESGVSVRKINVSPTTYRLKDCGFVTPRSRELPNPNSEILPGICHTDAY 660  
 601 VTLANKFESGVSVRKINVSPTTYRLKDCGFVTPRSRELPNPNSEILPGICHTDAY 660  
 DB 601 VTLANKFESGVSVRKINVSPTTYRLKDCGFVTPRSRELPNPNSEILPGICHTDAY 660  
 QY 661 QOGTNSKPCFOKKDVEDGNANFLGKASGID 690  
 661 QOGTNSKPCFOKKDVEDGNANFLGKASGID 690  
 DB 661 QOGTNSKPCFOKKDVEDGNANFLGKASGID 690

DB 661 QOGTNSKPCFOKKDVEDGNANFLGKASGID 690  
 RESULT 2  
 CYG3\_RAT  
 ID CYG3\_RAT STANDARD; PRT; 690 AA.  
 AC P19686;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanylate cyclase soluble, alpha-1 chain (EC 4.6.1.2) (GCS-alpha-1)  
 DE (soluble guanylate cyclase large subunit).  
 GN GUCY1A1 OR GUCY1A.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID:10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Lung;  
 RX MEDLINE=91009100; Pubmed=1698769;  
 RA Nakane M., Arai K., Saheti S., Kuno T., Buechler W., Murad F.;  
 RT "Molecular cloning and expression of cDNAs coding for soluble  
 guanylate cyclase from rat lung.";  
 RL J. Biol. Chem. 265:16841-16845(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Striatum;  
 RX MEDLINE=97151525; Pubmed=8997507;  
 RA Smigrodzki R.M., Levitt P.;  
 RT "The alpha 1 subunit of soluble guanylyl cyclase is expressed  
 prenatally in the rat brain.";  
 RL Brain Res. Dev. Brain Res. 97:226-234(1996).  
 CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.  
 CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF  
 MAGNESIUM OR MANGANESE IONS.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE  
 FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.  
 CC -1- SIMILARITY: BELONGS TO ADENYL CYCLASE CLASS-4/GUANYLYL CYCLASE  
 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M57405; AAA41206.1; -  
 DR EMBL: U60835; AAB17953.1; -  
 DR PIR: A38297; OYRTAL.  
 DR HSP: P19687; 1AMN.  
 DR Interpro: IPR001054; G\_cyclase.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART: SM00044; CYCC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS0125; GUANYLATE\_CYCLASES\_2; 1.  
 KM Lyase; GMP synthesis; Multigene family.  
 FT DOMAIN 480 607 GUANYLATE CYCLASE.  
 SQ SEQUENCE 690 AA; 77566 MW; E4819B2CA4F86401 CRC64;  
 Query Match 90.0%; Score 3232; DB 1; Length 690;  
 Best Local Similarity 89.1%; Pred. No. 4.3e-218;  
 Matches 616; Conservative 39; Mismatches 34; Indels 2; Gaps 2;

QY 1 MFCSTKLKDKITGECFSLAPGQVNESEEAAGSSSEKATVPICODIPEKNIOESLP 60  
 1 MFCSTKLKDKITGECFSLAPGQVNESEEAAGSSSEKATVPICODIPEKNIOESLP 60  
 DB 1 MFCSTKLKDKITGECFSLAPGQVNESEEAAGSSSEKATVPICODIPEKNIOESLP 60  
 QY 61 QRTSRSRYLHTLAESICKLIPPEERLNVALLQRTLAKHKIKESKSLEREDFEKTIAE 120  
 61 QRTSRSRYLHTLAESICKLIPPEERLNVALLQRTLAKHKIKESKSLEREDFEKTIAE 120  
 DB 61 QRTSRSRYLHTLAESICKLIPPEERLNVALLQRTLAKHKIKESKSLEREDFEKTIAE 120

[illegible]

```

RT "Catalytic mechanism of the adenyllyl and guanylyl cyclases: modeling
RL and mutational analysis.",
CC Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419(1997).
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS.
CC -1- SUBUNIT: HETEROIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X54014; CA37960.1;
DR PIR, S10713; OYB07.
DR PDB; 1AWN; 28-JAN-98.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYC; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.
DR Lysase; GMP synthetase; Multigene family; 3D-structure.
DR KW DOMAIN 482 609 GUANYLATE CYCLASE.
DR SEQUENCE 691 AA; 77532 MW; 5D1FEAD2204E8683 CRC64;
SQ
Query Match 86.5%; Score 3107.5; DB 1; Length 691;
Best Local Similarity 86.1%; Pred. No. 21e-209;
Matches 595; Conservative 41; Mismatches 54; Indels 1; Gaps 1;:
OY 1 MCTYKLDKLTITGECFSLAPGQVYNSSSEAAAGSSSESCATYVP-ICQDIPERNKIQSL 59
DB 1 MFCATKQLQITDPCFSLAPGQVRRPLCEATGSCGAPSPGCGVGPVDPKMPRL 60
OY 60 PORTSRBYVLTIAESICKLIPFEERLVAALQRTAKIKRKISRSLEDEPEKITA 119
DB 61 PRKTSRBYVLTIAESICKLIPFEERLVAALQRTAKIKRKISRSLEDEPEKITA 120
OY 120 EQAAVAGPVVEVYKESLGEVEFKICEEDENILGVAGGTLLKDLFSLFKOSSHCOEA 179
DB 121 DQALTAGPVPEIKESLGEVEFKICEEDEVYLGAVGGTLDLFLNSFTLLKQSHCOEA 180
OY 180 GKRGHLEVASILCDKEDDELHYVYFFPKRTSLIPGIKAAAHVLYTEVEVSLMPC 239
DB 181 EKKGFEDVASILCDKEDDVLVYVYFFPKRTSLIPGIKAAAHVLYTEVEVSLMPC 240
OY 240 FHNDCSEVNPQYLLSYTHMKSTRKSLSPKPOSSLVPTSLFCFEPFHFMDMTIL 299
DB 241 FHQDCREVVDQPCELSYHHSARHPPGKPVSSLVTPASLFCFEPFHFMDMTIL 300
OY 300 QFGNGIRLLMNRDFOGKPNFEFEFELLTPKINQTFSGIMTLMNQFVVRMRMDSYK 359
DB 301 QLGHGIRLLMNRDVOGKPNFEFEFELLTPKISQFSGIMTLMNQFVVRMRMDSMK 360
OY 360 SSRVMDLKGOMIYIYESSAILFLGSPCYDLDEDFGRGLYSDDIPINHALDVTILG 419
DB 361 SSRVMDLKGOMIYIYESSAILFLGSPCYDLDEDFGRGLYSDDIPINHALDVTILG 420
OY 420 RAODGLKRLRLKATLEQAHQALEEKKRKIVDILCSIFPEVAAQQLQGVQVQAKRFSN 479
DB 421 RAODGLKRLRLKATLEQAHQALEEKKRKIVDILCSIFPEVAAQQLQGVQVQAKR 480
OY 480 VTMLEFSDIVGTAICSCSPQVITMLNALTTRPQOCCGELDYKVEITIGAYCYAGGL 539
DB 481 VTMLEFSDIVGTAICSCSPQVITMLNALTTRPQOCCGELDYKVEITIGAYCYAGGL 540

```

```

QY 540 KESTHAVOIALMALKMEISDEVMSPHGEPIKRIKIGHSGSVAGVGVKMPRYCLFGN 599
DB 541 KESTHAVOIALMALKMEISHEVYSPGEPICKRIGHSGSVAGVGVKMPRYCLFGN 600
QY 600 NVTLANFESSGVPRKINVSPTTYRLKDCGFFVTPSPRELPNPNSEIDGICHPEDA 659
DB 601 NVTLANFESSGVPRKINVSPTTYRLKDCGFFVTPSPRELPNPNSEIDGICHPEDA 660
QY 660 YQOGTNSKPCPKKVEDGDNANFLGKASGID 690
DB 661 YQOGTNSKPCPKKVEDGDNANFLGKASGID 691

RESULT 4
CYGA_RAT CYGA_RAT STANDARD; PRT; 730 AA.
AC 09WV14;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanylate cyclase soluble, alpha-2 chain (EC 4.6.1.2) (GCS-alpha-2).
GN GUCYL12.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistar Kyoto; TISSUE=Aorta;
RX MEDLINE=20571097; PubMed=11121588;
RA Koglin M., Behrends S.;
RT *Cloning and functional expression of the rat alpha(2) subunit of
RT soluble guanylyl cyclase.;
RL Biochim. Biophys. Acta 1494:286-289(2000).
CC -1- FUNCTION: HAS GUANYLYL CYCLASE ON BINDING TO THE BETA-1 SUBUNIT.
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR HSMPL; AF109963; AAD42949.2; -.
DR HSPB; P19687; IAWN.
DR InterPro: IPR001054; G_cyclase.
DR Pfam: PF00211; guanylate_cyc; 1.
DR SMART: SM00044; CYC; 1.
DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE: PS01025; GUANYLATE_CYCLASES_2; 1.
DR Lysase: CGMP synthesis; Multigene family.
FT DOMAIN 519 646 GUANYLATE CYCLASE.
FT DOMAIN 50 72 ALA-RICH.
FT DOMAIN 64 72 POLY-ALA.
SQ SEQUENCE 730 AA; 81786 MW; B8D790BF81FB8F9 CRC64;

Query Match 46.0%; Score 1651; DB 1; Length 730;
Best Local Similarity 49.2%; Pred. NO. 1.le-107;
Matches 355; Conservative 106; Mismatches 196; Indels 64; Gaps 14;

QY 13 GECPEFSILA-----PGQVNPSSSEBAASSESCKATVPICODIPKNIQESL----- 59
DB 27 GECPEFSILCNNGSGSPG-----PPSGRAAMAMATP-----VPAASVAAAAAAYAG 73

```

```

QY 60 POKTSRBYVLAHTLAESICKLIFEPERENVALORTLA--KHKI-----KESRKL----- 109
DB 74 SKRAQRNRVRLDLSGESISLTLFAPSPQTHMLTKRLQYEHQVIGYRQAKRNFHNS 133
QY 110 -----ERDFEKTIAEQVAAGVPEVATKESIGEVEFKICYEDENILGVGCT 158
DB 134 RCSSADHSNKEIEIDVSGILNCTANVLGKQEOIERGEFFKICDENRVLAVGST 193
QY 159 LKDFLNSPTLLKSSHQEA--GKRGRLDASITCLKEDFLVYVFFPRTSLILPG 217
DB 194 LQDFNGFDALLE--HRTSPGKQATLESFICKELPSTGLTLNHFPHHTVGPALG 250
QY 218 IIRAAHVLYETEVEV--SLMPCEPHNDCSEFVNOPYLLVYVHMSKTPSLSPKQSSLY 276
DB 251 MIKAAGRIYHLNVEVEQIEENKFCSDGSPSNVSCFLFLKEKET--TQITKNIPQSTQ 309
QY 277 IPTSL-----FCGTPPHFDFDKMTILLQGNQIRLRMRDQGRNFEETLTPK 330
DB 310 IPTDLRISINTFCFTFFFLMFDPNMVVLQLEGRLQOL--RCDNHKYLKEDCEIYSPK 368
QY 331 INTFSGIMTLMNQFVVRVRMRDMSYKSSRVADLKGOMTYIYESSALIFLGSQVRL 390
DB 369 VNAIFDRLYLRLSTPFYIRKPEASGDNEDKVEIKGOMIHVESNAILFLGSPCYDKL 428
QY 391 EDFTGRGLYSDIPDIHNAIRDVYLIGQARAQDLKRLKATLQAHQALEEKKKT 450
DB 429 DELIGRGLHSDIPDIHNAIRDVYLIGQARAQDLKRLKATLQAHQALEEKKKT 488
QY 451 VDLICSTFPCVAAQMLQGVQAKKNSNTYMLFSDYVGFALTSQSPLOYITMLNLY 510
DB 489 VDLISITPCVAAQMLQGVQAKKNSNTYMLFSDYVGFALTSQSPLOYITMLNLY 548
QY 511 TRFDQCGELDYVYETIGDAYCVAGGLHKSDFHVAOIALMALKMEISDEVMSPHGE 570
DB 549 TRFHQCGFLDIYVETIGDAYCVAGGLHKSDFHVAOIALMALKMEISDEVMSPHGE 608
QY 571 IKMIGHSGSVPRKINVSPTTYRLKDCP 630
DB 609 IOMKIGHSGSVLGVVGRMPRYCLGNNTLASKESGSHPRINISPTTYQLKRED 668
QY 631 GFVTPSPRELPNPNSEIDGICHPEDAVQOGGINSRCPQKKNVEGNNAN----FLGKA 686
DB 669 SFTFIPSPRELPNPNSEIDGICHPEDAVQOGGINSRCPQKKNVEGNNAN----FLGKA 728
QY 687 S 687
DB 729 S 729

RESULT 5
CYGA_HUMAN CYGA_HUMAN STANDARD; PRT; 732 AA.
AC P33402;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanylate cyclase soluble, alpha-2 chain (EC 4.6.1.2) (GCS-alpha-2).
GN GUCYL12 OR GUCI12 OR GUCSA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92070494; PubMed=1683630;
RX Hatteneck C., Medel B., Koesling D., Malkewitz J., Boehme E.,
RA Schultz G.;
RT "Molecular cloning and expression of a new alpha-subunit of soluble
RT guanylyl cyclase. Interchangeability of the alpha-subunits of the
RT enzyme.";
FT FEBS Lett. 292:217-222(1991).
RN [2]

```

```

RP SEQUENCE FROM N.A. (ISOFORM ALPHA 2-1).
RA MEDLINE-95403342; PubMed-7673142;
RA Behrends S., Harteneck C., Schults G., Koesling D.:
RA "A variant of the alpha 2 subunit of soluble guanylyl cyclase
RT contains an insert homologous to a region within adenylyl cyclases
RT and functions as a dominant negative protein."
RL J. Biol. Chem. 270:21109-21113(1995).
CC -1- FUNCTION: HAS GUANYLYL CYCLASE ON BINDING TO THE BETA-1 SUBUNIT.
CC THE ALTERNATIVE SPLICED ISOFORM, ALPHA-2-1, ACTS AS A NEGATIVE
CC REGULATOR OF GUANYLYL CYCLASE ACTIVITY AS IT FORMS NON-FUNCTIONAL
CC HETERODIMERS WITH THE BETA SUBUNITS.
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA-2 (SHOWN HERE) AND
CC ALPHA-2-1. ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: ALPHA-2 ISOFORM IS EXPRESSED IN FETAL BRAIN,
CC LIVER, COLON, ENDOTHELIUM AND TESTIS, WHILE THE ALPHA-2-1 FORM
CC IS EXPRESSED ONLY IN LIVER, COLON AND ENDOTHELIUM.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLYL CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X63282; CAA44921.1; -
DR EMBL: Z50053; CAA90393.1; -
DR PIR: S18325; S18325.
DR HSRP: F19687; IAWN.
DR Genew: HGNC:4684; GUCY1A2.
DR MIM: 601244; -
DR InterPro: IPR001054; G_cyclase.
DR Pfam: PF00211; guanylate_cyc; 1.
DR SMART: SM00044; CYCC; 1.
DR PROSITE: PS00432; GUANYLYL CYCLASES_1; 1.
DR PROSITE: PS0125; GUANYLYL CYCLASES_2; 1.
DR KMW: CGMP synthesis; Multigene family; Alternative splicing.
FT DOMAIN 521 648 GUANYLYL CYCLASE.
FT DOMAIN 51 76 ALA-RICH.
FT VANSPLC 612 612 POLY-ALA.
FT VANSPLC 612 612 Q -> QPQRSEILSFVSIQIVPDQHOSETDUGTER
SQ SEQUENCE 732 AA; 81749 MW; 79465A7D3FE52DB7 CRC64;
Query Match 45.5%; Score 1635; DR 1; Length 732;
Best Local Similarity 49.4%; Pred. No. 1.5e-106;
Matches 355; Conservative 106; Mismatches 200; Indels 58; Gaps 15;
OY 13 GECFSLIA-----PGQVNESEEAAGSESCAKVPLICQDIPKNIKIGSLPORTS 65
DB 27 GECPLSLRCWNGSRSPG--PLESPAAAAA--AAPPTPAAAAAATAAGARROR 81
OY 66 RSRVYLTLASICKLIPFERERLVALQRTLA--KHKI--KESKRSL----- 109
DB 82 RRRVYLDISGISISRLAPSPQTIQTLKRTLOYEHOVIGYRDAEKNHNSNCVAD 141
OY 110 -----ERDFEKTAIEQAVAGVPEVYIKESIGEVEVICYEDENILGVAGTAKDFLN 164
DB 142 HSNKEIDVSGIIQCTANILGLKFEELQKRGEFFENICENRVLRAVAGTQADFFN 201
OY 165 FSFTLLKQSHQQA-GRGRLEDAISILCKEDKDFLHYVFFPKRTSLIPGITKAA 223
DB 202 GFDALLE---HRTSFGKATLESFSLCKELPECTIMLHTRPHHYIGFAMIGIKAA 258

```

```

OY 224 HLYETEVEVSLMPCCFEND--CSEFVNQ---YLYSVHMKSTKPSLSPKPOSSIVP 278
DB 259 KRIYLDVEBQVA-----NEKICSD-VSNPGNCSCLFTFLICECENTIMKMLPGTQSPV 313
OY 279 TSL-----FCKTFPHFMPKDMTILDPGKIRRLAMRDPQCPNPEEFELITPKIN 332
DB 314 ADLRISINTFCRAFPFHFMLFSPMSVQLDGLRLKQI--RCDTHKVLKPEDCFEIVSPKVN 372
OY 333 QFSGIMTMNLNQFVVRVRMDNSYKSSRVWDKGMQIYVESAILFISPCVDRD 392
DB 373 ATFEVRLRLSTPFIYKIRKPPASGSENDKMEVAGQIHVPESISILFSLSPCYDKDE 432
OY 393 FTGRGLYLSDIPINHALDVLVIGBQARAQDLKRLKRLKATLEBOAQALEEKRTVD 452
DB 433 IMGRHLHSDIPIDHATDVLVIGBQARAQDLKRLKRLKATLEBOAQALEEKRTVD 492
OY 433 ILCSIFPEVNAQOLMGQVQAKRSNTMLFSDIVGTALCSQSPLOVITMLALYTR 512
DB 493 LLYSIFPDVDAQOLMGQVQAKRPDVTMLFSDIVGTALCAOCTPMQVISMNELYTR 552
OY 513 FDOQCELDVYKVENTIGDAYCVAAGLHRESPTHAQVIALMKMELSDVWSPHGEPIK 572
DB 553 FDHOGFLDIYKVENTIGDAYCVAAGLHRSKLSCHAPRIALMKMELSEVILTPGRIQ 612
OY 573 KRIGHSQSVAGVGVKMPRYCLEGNNTVLANKFESQSVPRKINVSPTTYRLKDCPGF 632
DB 613 KRIGHSQSVYAGVGVKMPRYCLEGNNTVLANKFESQSHPRINVSPTTYRLKRESF 672
OY 633 VTPRSRELPNPFSEIRPGICFLDAVOQGNNSKPCQKQKDYEGNAN----PLGRKS 687
DB 673 TFIPRSRELPNPFSEIRPGICFLVETRGPRKPPSLSSRIKRVSNIGTMFLRETS 731

```

## RESULT 6

```

CGH_DROME STANDARD; PRT: 683 AA.
ID CGH_DROME
AC 007093;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Head-specific guanylate cyclase (BC 4.6.1.2).
GN GYC-ALPHA-99B OR GYC-ALPHA-63A OR GYC OR DGC1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Canton-S; TISSUE=Head;
RX MEDLINE-93203896; PubMed-8095978;
RA Yoshikawa S., Miyamoto I., Aruga J., Furutachi T., Okano H.,
RA Mikoshiba K.,
RT "Isolation of a Drosophila gene encoding a head-specific guanylyl
RT cyclase."
RT J. Neurochem. 60:1570-1573(1993).
CC -1- FUNCTION: MAY HAVE A ROLE IN PHOTOTRANSDUCTION. A SECOND SUBUNIT
CC MAY BE REQUIRED FOR ENZYME ACTIVITY.
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- SUBUNIT: DIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- TISSUE SPECIFICITY: HEAD, WHERE IT IS PREFERENTIALLY EXPRESSED IN
CC THE CNS AND THE RETINA. NOT FOUND IN BOOIES.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: S57126; AAB25820.1; -.
DR PIR: JH0810; JH0810.
DR HSSP: P19687; 1AMN.
DR FLYBase: FBgn0013972; Gyc-alpha-99B.
DR InterPro: IPR001054; G_cyclase.
DR Pfam: PF00211; guanylate_cyc; 1.
DR SMART: SM00044; CYC; 1.
DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; FALSE_NEG.
DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 1.
DR Lysase: CGMP synthesis; Multigene family; Vasion.
FT DOMAIN 465 591 GUANYLATE CYCLASE.
SQ SEQUENCE 683 AA; 75906 MW; 12D67C4B79A07C8D CRC64;

Query Match 27.2%; Score 976.5; DB 1; Length 683;
Best Local Similarity 37.3%; Pred. No. 1.3e-60;
Matches 244; Conservative 116; Mismatches 211; Indels 83; Gaps 20;

CC -----
DB 58 SLPOKRSRRLVLTALAEISCKLIFEEFERLNVALLORTLAK-----HKIK---ESR 106
DB 27 ALDEEDLSDATLTTLQMAIQLTAPSNEDLNAVSLVAKYROMPNHKLKLDPT 86
CC -----
QY 107 KSLREDEFEKTIAE-----QAVAGVPEVIEKESLGEVEFKICYEE-DENILGVG 159
DB 87 KSCANDYIADIDQELLKMDASSELV-----LLGEELITCCCTGIIERARCLG 141
QY 160 KDFLNSFS---TLIKSSHCOEAGKRGRLDASTLCKEDDLHYFFFPKRTSLIP 216
DB 142 QEEFLSGIDGYDVLYK-----LQEEED---VTDGFCVAGEGE---LIFTSE 189
QY 217 GIILKAHVLYETEVESLMPCEHNDSEFVNQPYLL-----YSVHKSIRKPSLSP 272
DB 190 GSILALTRMLYKVDVNTKIEPV---ERCPAVSLPLLAGOGQLPDHAGSSTSVK 245
QY 273 -----SSVYIPSTLFCCTEPEFHEFMDKMTILQFNGIRLNRDRFOGKPF 322
DB 246 TVQRNSNSNSADLOMNSSSCCKMPHMFIMNEOLEVQLRGSKLY-----K 298
QY 323 -----YEELIPK-INQTFSGIMTLMNQFVVRVRMRDMSYKSSRVNLDGOM 374
DB 299 FGCAATTFYFPRKPKGLTMKFRDIVRTYPTFLGLNPNPGAVDFPAIGLEIK 358
QY 375 ESSAILEFSPCYDRLEDFTGRGLYLSIDIPINHALRNVLYLGGQARODGLKRL 434
DB 359 ESSNLSLTFGSPFDGLDGLTCNGFLISDIPLDHATREVLIVGQARODGLRRM 418
QY 435 TLBOAHQALEEKKKTYDLCSIFPCVAAQLMGOVVOAKKFSNVTMLFSDIYGA 494
DB 419 SIEANSAYTKERKKNSLHLIFPAELAKMLGSSIDAKTYPDVTIILFSDIVG 478
QY 495 SQCSPLQVITMLNALYTRPDQCGELDYVETIGDAYCVAGLHKS--DTNAVOT 552
DB 479 SRATPFVYISMLEGKDFDEFCDFEFVYKETIGDAYCVASGLHRASTIDAR 535
QY 553 ALKAMELSDVMSPHGEPKIKRIGLSGSAFAGVVGKMPRYCLFGNNVTLAKRES 612
DB 536 GLKMLDSCSHIHHDGQIKRIGLHAGIVLAGVGGKMPRYCLFGNSVITANKF 595
QY 613 PREINVSPTTYRLKDCQGFV--TPRSRELPNPSESLPG---ICHFLDVO 661
DB 596 ALKINVSPTTKDMLTKHEGFEEFELQPRDPSFLPKPEPN--PGGFEICYF 647

RESULT 7
CYG1_RAT
AC P30595;
ID CYG1_RAT STANDARD; PRT; 619 AA.
DT 01-FEB-1991 (Rel: 17, Created)
DT 01-FEB-1991 (Rel: 17, Last sequence update)
DT 16-OCT-2001 (Rel: 40, Last annotation update)
DE Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (GCS-beta-1)
DE (Soluble guanylate cyclase small subunit).
GN GUCY1B1 OR GUCY1B3 OR GUC1B3.

```

```

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89087429; PubMed=2905128;
RA Nakane M., Saheki S., Kuno T., Ishii K., Murad F.;
RT "Molecular cloning of a cDNA coding for 70 kilodalton subunit of
RT soluble guanylate cyclase from rat lung."
RL Biochem. Biophys. Res. Commun. 157:1139-1147(1988).
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: LUNG AND BRAIN.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25562; AAA41204.1; -.
DR PIR: A31871; OYRBL.
DR HSSP: P16068; 1AMN.
DR InterPro: IPR001054; G_cyclase.
DR Pfam: PF00211; guanylate_cyc; 1.
DR SMART: SM00044; CYC; 1.
DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 1.
DR Lysase: CGMP synthesis.
FT DOMAIN 421 554 GUANYLATE CYCLASE.
SQ SEQUENCE 619 AA; 70477 MW; BCG9850DE4504226 CRC64;

Query Match 22.1%; Score 795; DB 1; Length 619;
Best Local Similarity 34.3%; Pred. No. 5.2e-48;
Matches 210; Conservative 100; Mismatches 225; Indels 78; Gaps 15;

CC -----
QY 89 LNVALLORTLAKHKIKESKSLERE-----DEKT---IAEQVAAGVPE 130
DB 5 VNALELVIRNNGPREWEDIKKEADLODEGQPLVITIIDHSTYDLVLAASVNLNAG 64
QY 131 VIKESLGEVEFKICYED-ENLGVVGTLLKDFLNSFTLLKSSHCOEAGKRGRLDAS 189
DB 65 EILQMGKMFVFCQSGVDTLIRLVGSNREFLQMLDL-----HDHLATYIPGRAPS 119
QY 190 IICLDNE-DDFLAVYFFFPKRTSLILPGIKRAAHVLYETEVESLMA---PPCHNDC 244
DB 120 FRCYDKEKRGKGLIHYHSEREGIADIVIGIKTYVAQIIGETELDKMVFVIOORSE 178
QY 245 SEFVNQPYLLYSVHMKSTPS-----LSPKPOSSVYIPSTLFCCTEPEFHEFMD 296
DB 179 -----FLIEEKESKEEDYEDLDLFEENGQDSKISPTT-TCAPFPPIIITDR 227
QY 297 TIIQFNGIRLRNLRDFOGKPFEEYFELLPKINQTFSGIMTLMNQFVVRVRMRDMS 356
DB 228 VVQCGNAIYRVLPQI-PGKSLSVFSVLRHIDISFHGLISHINTVFVLRSKEDGLD 286
QY 357 VKK-----SSRVNLDGOMITIVSSALFLGSCVDRLDFTGRGLYSLDIP 406
DB 287 VEKLECEDELTAIEISCLRLKRGKATLEQAHQALEEKKKTYDLCSIFPCVAAQL 466
QY 407 NMLRDVVLIGEORADGKRRGRKATLEQAHQALEEKKKTYDLCSIFPCVAAQL 466
DB 347 DATRDVLVGEORPEYRKLTQLELITLDRQLTLRALDEKKKTYDLCSIFVLPVANEL 406

```





```

RN      FEBS Lett. 304:83-88(1992).
RP      [2]
RA      SEQUENCE FROM N.A. (ISOFORM HSGC-2).
RB      TISSUE-Kidney;
RC      Ganssems Y., Brouckaert P., Fiers W.;
RD      Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
RE      [3]
RF      RN
RG      TISSUE-Of 337-545, AND ALTERNATIVE SPLICING.
RH      MEDLINE-92008652; PubMed-1680753;
RI      Chhajlani V., Fraendberg P.-A., Ahlner J., Axelsson K.L.,
RJ      Wikberg J.E.S.;
RK      "Heterogeneity in human soluble guanylate cyclase due to alternative
RL      splicing.";
RM      FEBS Lett. 290:157-158(1991).
RN      CC
RP      CC -1- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.
RC      CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
RD      CC MAGNESIUM OR MANGANESE IONS.
RE      CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
RF      CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
RG      CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: HSGC-1 (SHOWN HERE) AND HSGC-2;
RH      CC ARE PRODUCED BY ALTERNATIVE SPLICING.
RI      CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
RJ      CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
RK      CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
RL      CC FAMILY.
RN      CC
RP      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RC      CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RD      CC the European Bioinformatics Institute. There are no restrictions on its
RE      CC use by non-profit institutions as long as its content is in no way
RF      CC modified and this statement is not removed. Usage by and for commercial
RG      CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
RH      CC or send an email to license@isb-sib.ch).
RI      CC
RJ      CC EMBL; X66533; CAA47144.1;
RK      CC EMBL; AF020340; AAB94877.1;
RL      CC PIR; S23097; S23097.
RN      DR
RP      DR HSSP; P16068; 1A9N.
RC      DR Genew; HGNC:4687; GUCY1B3.
RD      DR MIM; 139397;
RE      DR InterPro; IPR001054; G-Cyclase.
RF      DR Pfam; PF00211; guanylate_cyc; 1.
RG      DR SMART; SMO0044; CYC; 1.
RH      DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
RI      DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.
RJ      DR Lyase; cGMP synthetis; Alternative splicing.
RK      FT DOMAIN 421 554 GUANYLATE CYCLASE.
RL      FT VASPLIC 393 425 MISSING (IN ISOFORM HSGC-2).
RN      SO
RP      SO SEQUENCE 619 AA; 70514 MW; 21EAE66DE02A1 CMC64;
RN      Query Match 22.0%; Score 789.5; DB 1; Length 619;
RP      Best Local Similarity 35.1%; Pred. No. 1.3e-47;
RD      Matches 207; Conservative 96; Mismatches 219; Indels 67; Gaps 15;
RN      Db
RP      Db 101 KIKRSKSLSEDEDEKTIIEAQAAGPV-----EYIKSLGEYVRICTEED-ENILG 153
RC      Db 30 QLDSEGQFLVNIITDDSKTYDVAASAQVLMNAGELL-QMFGMFVFCQESGYDILR 88
RD      Db 154 VGGTILDFLNSFTLLKQSSHCQAGKRGRLDASTILCKDE-DDFLHYVFFPKRTTS 212
RE      Db 89 VLGSNVAEFLQNDAL-----HDLIATITPQMRAPSRCTDAEKGKGLILHYISERGLO 143
RF      Db 213 LLPIGIIKAAAHVLYETEVEVSLM-----PQCFHNDCESEFVNQPYLLYSVHKSTKPS-- 265
RG      Db 144 DIVIGIIRKVAQOIHGEIEMDKVLIQGNNECDHTQ-----FLIEKEKEDDFY 192
RH      Db 266 -----LSPSKQSSLVIPISLFCETFPFHMEKDKMTILFQNGIRRLMNRDFQGRPNF 320
RI      Db 193 EDDLDFRENGEQESRISPYT-FCRAKPFPHIIFRDLVAVTCGNAIYVTLQI-QGNCSTL 250
RJ      Db 321 EYVEYELLPRKNQFSGIMTLNQGQVAVRVRARMNSYRK-----SSRVMDLKGOM 370

```

ID	CYC2_RAT	STANDARD	PRT:	682 AA.
Db	251	LSVSLVPHIDISPHGLISHINIFVLRSKEGLLDVEKLBCDEBLPTETLSCLRLKQGM	310	
Qy	371	IYIVESSAILFLGSPCDRLDEDFRGGLYSDIPFHNALRQVYLIGEDARADGLKRLG	430	
Db	311	IIYFPEADSIILFLGSPSVNNLDLTRRGLYLDIPLDATRDLVLLGGEFFREYKLTQELE	370	
Qy	431	KLKATLEAOHAALEEKKTVDLCSIPCEVAOOLMGQGVVQAKKFSNVMMLPESDIYGF	490	
Db	371	ILTRDLQTLTRALDEDEKKTTLTXSVLPSPVANLHRKRKRPAPKADYNVITLISGVIF	430	
Qy	491	TAICSQCS---PLQVITMLNALYTRPDQCGELD---VYAEVETIGDAYCYAGGLHKESED	543	
Db	431	NAFCSKHAHGSGAKKIYNALNDLITRFTPLTDSRKNRPVYIVETVGDYKMTVSGLEPCI	490	
Qy	544	THAVOIALMALKMEIODEVMSPHGEPKIRKIGHSGSVFAGVGVKMPRYCLFEGNVTL	603	
Db	491	HAASICHIALDMMHIAQGV---QVGSSEVQITIGHTGEVYGVIGQRPBRYCLFEGNVTL	549	
Qy	604	ANKFESCVPRKINVSPPTVRYLLDCCPEFVTPPSREELPNPFPSEIG	652	
Db	550	TSRTEITGEGKGINVSEYTYNCL-----MSPENDSQFLEHRG	588	
RESULT 10				
ID	CYC2_RAT	STANDARD	PRT:	682 AA.
Ac	P22717			
Dt	01-AUG-1991	(Rel. 19, Created)		
Dt	01-AUG-1991	(Rel. 19, Last sequence update)		
Dt	16-OCT-2001	(Rel. 40, Last annotation update)		
De	Guanylate cyclase soluble, beta-2 chain (EC 4.6.1.2) (GCS-beta-2).			
Gn	GUCCY182.			
Os	Rattus norvegicus (Rat).			
Oc	Euarctota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Oc	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
Ox	NCBI_Taxid:10116;			
Rn	[1]			
Rp	SEQUENCE FROM N.A.			
Rc	TISSUE=Kidney;			
Rx	MEDLINE=91105012; PubMed=1980215;			
Rt	Yuen P.S.T., Potter L.R., Gathers D.L.;			
Rt	*A new form of guanylyl cyclase is preferentially expressed in rat			
Rt	kidney.*;			
Rl	Biochemistry 29:10872-10878 (1990)			
Cc	-1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.			
Cc	-1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF			
Cc	MAGNESIUM OR MANGANESE IONS.			
Cc	-1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.			
Cc	-1- SUBCELLULAR LOCATION: Cytoplasmic.			
Cc	-1- TISSUE SPECIFICITY: KIDNEY AND LIVER.			
Cc	-1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE			
Cc	FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.			
Cc	-1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE			
Cc	FAMILY.			
Cc	-----			
Cc	This SWISS-PROT entry is copyright. It is produced through a collaboration			
Cc	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
Cc	the European Bioinformatics Institute. There are no restrictions on its			
Cc	use by non-profit institutions as long as its content is in no way			
Cc	modified and this statement is not removed. Usage by and for commercial			
Cc	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
Cc	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
Cc	-----			
Dr	EMBL; M57507; AAA41207.1; -			
Dr	PIR; A36228; OYRTB2.			
Dr	HSP; P19687; 1AMN.			
Dr	InterPro; IPR001034; G_cyclase.			
Dr	InterPro; IPR001230; Prenyl_site.			
Dr	Pfam; PF00211; guanylate_cyc; 1.			
Dr	SMART; SM00044; CYC; 1.			
Dr	PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.			
Dr	PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.			
Dr	Protein synthesis.			



```

FT DOMAIN 408 536 GUANYLATE CYCLASE.
SQ SEQUENCE 682 AA; 76196 MW; 98C173C1ACCT715 CRC64;
Query Match 21.6%; Score 774.5; DB 1; Length 682;
Best local similarity 32.6%; Pred. No. 1.6e-46;
Matches 190; Conservative 107; Mismatches 204; Indels 81; Gaps 10;
OY 129 VEVIKESIGEVEFKICYEED-ENIGVVGTLKDFLNSFTLLKQSHQOENKGRRLD 187
DB 1 MEALIKLGEYEFKCKSGYDRMLRTLGNGLTERIENDL-----HSYALASYENMA 55
OY 188 ASILCLKEDDFLHYFFPKRTSLIPGIIKAAAHVLEYEVESLMPCHDCCSEF 247
DB 56 PSFVREBAGDAMLHYSDRGICHIYPGIIIEAVAKDFDIDVAMSLI-----DMNEE 109
OY 248 V-----NOPLYLTVHMKSTKP--SLSPSKQSS----- 274
DB 110 VERGKREHYVFLVYQAKHROIGAKASRPOGSESDQADQALQSTLKRERYINIPVC 169
OY 275 -----LVIPISLCKTPFPFHFDKMTILQFG 302
DB 170 PGEKSHYAVRASVYLFGRKPLDTPQPYPERLWVEEVEFCDAFPFHIFEDALRYKQAG 229
OY 303 NGIRLMMRRDQGRNFEVEFEITPKINOTFGSINTMLNQFVVRAR-WDNSYKSS 361
DB 230 VNIQKYPGILITQ-KFALDEYFSLIHPVTENISSICKFISQFVLKTRKEMPRARSQ 288
OY 362 RVMDLKGMIYIVESSALFSCVDRLEDFTRGLYSDIPINHALDVLIGEQA 421
DB 289 PMLKRGQIMNESLRCHMIFKSPVNRSTQLEESKMHSDIAPDTRDLILNQRLA 348
OY 422 QDGLKRLKLTALQEAHQALEEKKRTVDLCSIFPCVQAQOLMOGOVYQAKFSNVT 481
DB 349 EMELSCQLEKKEELRVLSNHLAIEKKRTETLLYMLPEHVANOQLEKGRKVAAGEFCT 408
OY 482 MLFSDIYGFALCSQSPLOVYTMALYTRDQCGGELDYKVTETIGDAYVAGLAK 541
DB 409 ILFSVVFPTNCAACEPIQIVNMLNSMSKPDRLTSVADYKVTETIGDAYVAGVGPV 468
OY 542 SDTHAVQIALMALKMMELSDVEVSP-HGEPIRMKIGLSGSAFVAGVGMRYCLFGNN 600
DB 469 VESHAQVRANFALGRISAKVEYMNVPYGEPIQIRGIRHGPVLAGVGDKMRVCLFGDT 528
OY 601 VTLANKFESCVPRKINVSPTRYLLKDCGFFVTPRSREL 642
DB 529 VNTASRMESHGIPSKVHLSPFAHRAKLN-KGFEIYRGEIEV 569

```

## RESULT 11

```

CYG2_HUMAN STANDARD; PRT; 617 AA.
AC 075343; G9NZ64;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN Guanylate cyclase soluble, beta-2 chain (EC 4.6.1.2) (GCS-beta-2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN 11
RP SEQUENCE FROM N.A., AND REVISION TO 175.
RC TISSUE=Kidney;
RX MEDLINE=20241821; PubMed=10777682;
RA Behrens S., Vehse K.;
RT "The beta(2) subunit of soluble guanylyl cyclase contains a
RL human-specific frameshift and is expressed in gastric carcinoma.";
RN Biochem. Biophys. Res. Commun. 271:64-69(2000).
RP SEQUENCE OF 95-257 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=99107820; PubMed=9889008;

```

```

RA Behrens S., Kazmierczak B., Steenpa A., Knaf B., Bullerdick J.,
RA Scholz H., Elberg H.;
RT Assignment of GUCY1B2, the gene coding for the beta2 subunit of human
RT guanylyl cyclase to chromosomal band 13q14.3 between markers D13S168
RT and D13S155.
RL Genomics 35:126-127(1999).
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
DR EMBL; AF038499; AAD09440.2; -.
DR EMBL; AF218383; AAF6105.1; -.
DR HSSP; P19687; 1ANN.
DR GENE; HGNC:4686; GUCY1B2.
DR MIM; 603695; -.
DR INTERPRO; IPR001054; G_Cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; FALSE_NEG.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
KM Lyase; CGMP synthesis.
FT DOMAIN 391 519
SQ SEQUENCE 617 AA; 70367 MW; 1B6C965CAB9A5E3E CRC64;
Query Match 19.7%; Score 708; DB 1; Length 617;
Best local similarity 31.3%; Pred. No. 6.1e-42;
Matches 179; Conservative 100; Mismatches 191; Indels 102; Gaps 12;
OY 149 ENILGVGGLKDFLNSFTL-----LKQSHQOENKGRRLDASTLCLDKEDD 198
DB 5 DMLRTLGNGLMEFTENDALHSYALASYENANSPFRERADGKM----- 51
OY 199 FLHYVFFPKRTSLILGIIKAAAHVLEYEVESLMPCHDCCSEVNP----- 251
DB 52 FLH--YSDRSGLCIHVPGIIIEAVAKDFDIDVIMDIL-----DMNEVEERTGKEHYV 103
OY 252 -YLTVSVH--MKSTKP----- 264
DB 104 FLIVQKARKMKRTPKRLQDSQGERQEAQAAFLKKEKYLIVNSACPYYKSHWDYR 163
OY 265 -----SLSPSKQSSLVIPISLCKTPFPFHFDKMTILQFGNGIRLMNR 312
DB 164 SYMFEGKHLMTFPIYPE-RLMIEKTFCAFPFHIFEDSLQKARVNIQYVPEL 222
OY 313 DQGRNFEVEFEITPKINOTFGSINTMLNQFVVRARMSYKSSR-VMDLKGMI 371
DB 223 QFO-NIQDEYFSLIHPVTENISIRRFISQFVLKTRKEMPRAMOSRTTLKIQGMI 281
OY 372 YIVESSALIFLSPCVDRLEDFTRGLYSDIPINHALDVLIGEQAQADGLKRLGK 431
DB 282 WMSNMCMWYLLCSPLKRSIQLELELNHMSDAPDTRDLILNQRLAIELEISNQLR 341
OY 432 LKATLEQAQALEEKKRTVDLCSIFPCVQAQOLMOGOVYQAKFSNVTMLFSDIYGF 491
DB 342 KKEELQVLSKHLAIEKKRTETLLYMLPEHVANOQLEKGRKVAAGEFSCITLFSVYET 401
OY 492 AICSQSPLOVYTMALYTRDQCGGELDYKVTETIGDAYVAGLHSESTHVAQAL 551
DB 402 NICTACEPIQIVNMLNSMSKPDRLTSVADYKVTETIGDAYVAGVGPVIGNHAQVAN 461

```

QY 552 MALKMELSDVNSP-HGEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNNVTLANKEESC 610  
 DB 462 FALGMRSIAKEVNPVYGEPIQLRWGIGHGTVLADVDGKMPRYCLFGDVTNTASRRESH 521  
 QY 611 SVPRKIVSPPTTYRLKDCPGFVFTPRSREL 642  
 DB 522 GLPKRVHLSPTAYRALKN-OGFKTIERGELEV 552

RESULT 12  
 ANPB\_BOVIN STANDARD; PRT; 1047 AA.  
 ID ANPB\_BOVIN STANDARD; PRT; 1047 AA.  
 AC P46197;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Atrial natriuretic peptide receptor B precursor (ANP-B) (ANPRB) (GC-B) (guanylate cyclase) (EC 4.6.1.2) (NPR-B) (Atrial natriuretic peptide B-type receptor).  
 GN NPR2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=95147865; Pubmed=7845391;  
 RA Fenrick R., Babinski K., McNICOLL N., Therrien N., Drouin J., de Lean A.,  
 RT "Cloning and functional expression of the bovine natriuretic peptide receptor-B (natriuretic factor R1c subtype).";  
 RL Mol. Cell. Biochem. 137:173-182(1994).  
 RN [2]  
 RP SOURCE OF 551-1047 FROM N.A.  
 RC TISSUE=Retina;  
 RX MEDLINE=93371497; Pubmed=8103329;  
 RA Simmlauer B.E., Zubov D.V., Abdulaev N.G.;  
 RT "Detection of expression of a membrane form of the guanylate cyclase type of GC-B in cattle retina (letter).";  
 RL Bioorg. Khim. 19:682-685(1993).  
 CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE CYCLASE ACTIVITY ON BINDING OF LIGAND. THE ACTIVATION ORDER SEEMS TO BE CNP > BNP > ANP.  
 CC -1- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C) WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.  
 CC -1- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC DOMAIN OF PROTEIN KINASES.  
 CC -1- SIMILARITY: BELONGS TO ADENYL CYCLASE CLASS-4/GUANYL CYCLASE FAMILY  
 CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L26359; AAC41619.1; -  
 CC EMBL: X66865; CAA47334.1; -  
 CC HSSP: Q02846; IAWL.  
 CC InterPro: IPR001828; ANP\_receptor.  
 CC InterPro: IPR000719; Euk\_kinase.  
 CC InterPro: IPR001054; G\_cyclase.  
 CC InterPro: IPR001170; Ntpep\_receptorin.

DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00211; guanylate\_cyc; 1.  
 DR Pfam; PF01094; ANP\_receptor; 1.  
 DR PRINTS; PR00255; NATPEPTIDER.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00044; CYC; 1.  
 DR PROSITE; PS00458; ANP\_RECEPTORS; 1.  
 DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;  
 KW cAMP\_synthetis; Signal.  
 FT SIGNAL; 1..22  
 FT CHAIN; 23..1047  
 FT DOMAIN; 23..458  
 FT TRANSMEM; 459..478  
 FT DOMAIN; 479..1047  
 FT DOMAIN; 513..786  
 FT DOMAIN; 861..991  
 FT DISULFID; 75..101  
 FT DISULFID; 439..439  
 FT DISULFID; 448..448  
 FT CARBOHYD; 24..35  
 FT CARBOHYD; 161..161  
 FT CARBOHYD; 195..195  
 FT CARBOHYD; 244..244  
 FT CARBOHYD; 349..349  
 FT CONFLICT; 958..958  
 SQ SEQUENCE 1047 AA; 117145 MW; 1A3814D0F22D64F CRC64;  
 Query Match 12.8%; Score 460; DB 1; Length 1047;  
 Best Local Similarity 41.9%; Pred. No. 2,6e-24;  
 Matches 113; Conservative 39; Mismatches 76; Indels 42; Gaps 6;  
 QY 411 DVLVIGEQARAOD-----GLKRLRG-----LKATLEDA 439  
 DB 760 EYLTLMEKRCWAODPAERPDGQIKGFIRRNKGGTSTIDNLLRMEQYANNLEKLYEER 819  
 QY 440 HQALEEKRRKTVLLCSIFCEVAQOOLMOGVQAKFSNVMTLFSDIYGFATICSQSP 499  
 DB 820 TOAYLEKKRAEALLVQILPHSAEOLKRGETVQALAFSVITYFSDIYGFATLSAESP 879  
 QY 500 LQVITMLNLYTRFDQCGELDLYKVTIGDAYCVAGL-HKESDTHAVQIALMAKME 558  
 DB 880 MQVVTLLNDLYTCFDALINDFDYKVTETIGDAYMVVSGLPGRNGQRHAEIARMALALL- 938  
 QY 559 LSEVWS-----PHEPIKMRIGLHSGSVFAGVGVKMPRYCLFGNNVTLANKEESC 612  
 DB 939 --DAVSSFRIHRPH-DQLRLRIGVHTGPACAGVGLKMPRYCLFGDVTNTASRRESNG 995  
 QY 613 PRKIVSPPTTYRLKDCPGFVFTPRSREL 642  
 DB 996 ALKIHVSSTTKDALDELGCPLERGDVEM 1025

RESULT 13  
 ANPB\_HUMAN STANDARD; PRT; 1047 AA.  
 ID ANPB\_HUMAN STANDARD; PRT; 1047 AA.  
 AC P20564; Q90050; O60871;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Atrial natriuretic peptide receptor B precursor (ANP-B) (ANPRB) (GC-B) (guanylate cyclase) (EC 4.6.1.2) (NPR-B) (Atrial natriuretic peptide B-type receptor).  
 GN NPR2 OR ANPRB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A. (LONG ISOFORM).  
RC TISSUE-BRAIN;  
RX MEDLINE-89365195; PubMed-2570358;  
RA Chang M.S., Lowe D.G., Lewis M., Hellmuss R., Chen E., Goeddel D.V.;  
RT "Differential activation by atrial and brain natriuretic peptides of  
RT two different receptor guanylate cyclases.";  
RL Nature 341:68-72(1989).  
RN [2]  
RP SEQUENCE FROM N.A. (LONG ISOFORM).  
RC TISSUE-Blood;  
RX MEDLINE-99182382; PubMed-10082481;  
RA Rehmanudla D., Nakayama T., Soma M., Takahashi Y., Uwabo J., Sato M.,  
RA Izumi Y., Kamatsuse K., Ozawa Y.;  
RT "Structure of the type B human natriuretic peptide receptor gene and  
RT association of a novel microsatellite polymorphism with essential  
RT hypertension.";  
RL Circ. Res. 84:605-610(1999).  
RN [3]  
RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
RC TISSUE-Kidney;  
RA Hirsch J.R., Meyer M., Magert H.J., Forssmann W.G., Mollerup S.,  
RA Herter P., Weber G., Cernak R., Ankortina-Stark I., Schlatter E.,  
RA Kruboffer M.;  
RT "GMP-dependent and independent inhibition of a K<sup>+</sup> conductance by  
RT natriuretic peptides. Molecular and functional studies in human  
RT proximal tubule cells.";  
RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE  
CC CYCLASE ACTIVITY ON BINDING OF ANP. SEEMS TO BE STIMULATED MORE  
CC EFFECTIVELY BY BRAIN NATRIURETIC PEPTIDE (BNP) THAN BY ANP.  
CC -1- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.  
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM/NPR-BI; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO  
CC WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)  
CC WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE  
CC CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.  
CC -1- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC  
CC DOMAIN OF PROTEIN KINASES.  
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE  
CC FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AB005647; BAA81737.1; JOINED.  
DR EMBL: AB005626; BAA81737.1; JOINED.  
DR EMBL: AB005627; BAA81737.1; JOINED.  
DR EMBL: AB005628; BAA81737.1; JOINED.  
DR EMBL: AB005629; BAA81737.1; JOINED.  
DR EMBL: AB005630; BAA81737.1; JOINED.  
DR EMBL: AB005631; BAA81737.1; JOINED.  
DR EMBL: AB005632; BAA81737.1; JOINED.  
DR EMBL: AB005633; BAA81737.1; JOINED.  
DR EMBL: AB005634; BAA81737.1; JOINED.  
DR EMBL: AB005635; BAA81737.1; JOINED.  
DR EMBL: AB005636; BAA81737.1; JOINED.  
DR EMBL: AB005637; BAA81737.1; JOINED.  
DR EMBL: AB005638; BAA81737.1; JOINED.  
DR EMBL: AB005639; BAA81737.1; JOINED.  
DR EMBL: AB005640; BAA81737.1; JOINED.  
DR EMBL: AB005641; BAA81737.1; JOINED.  
DR EMBL: AB005642; BAA81737.1; JOINED.  
DR EMBL: AB005643; BAA81737.1; JOINED.  
DR EMBL: AB005644; BAA81737.1; JOINED.  
DR EMBL: AB005645; BAA81737.1; JOINED.

DR EMBL: AB005646; BAA81737.1; JOINED.  
DR EMBL: AB005628; CA006466.1; -.  
DR PIR: S05514; S05514.  
DR HSSP: 002846; 1AWL.  
DR GeneW: HGNC:7944; NPR.  
DR MIM: 108961; -.  
DR InterPro: IPR001828; ANP\_receptor.  
DR InterPro: IPR00719; Euk\_pkinase.  
DR InterPro: IPR001054; G\_cyclase.  
DR InterPro: IPR001170; Ntpep\_receptor.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF00211; guanylate\_cyc; 1.  
DR Pfam: PF01094; ANP\_receptor; 1.  
DR PRINTS: PR00255; NATPEPTIDER.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00044; CYCC; 1.  
DR PROSITE: PS00458; ANP\_RECEPTORS; 1.  
DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
DR PROSITE: PS0125; GUANYLATE\_CYCLASES\_2; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;  
KW GMP synthesis; Signal; Alternative splicing; Polymorphism.  
FT CHAIN 1 22  
FT SIGNAL 1 22  
FT CHAIN 23 1047  
FT DOMAIN 23 458  
FT TRANSMEM 459 478  
FT DOMAIN 479 1047  
FT DOMAIN 513 786  
FT DOMAIN 861 991  
FT DISULFID 75 101  
FT DISULFID 439 439  
FT DISULFID 448 448  
FT CARBOHYD 24 24  
FT CARBOHYD 35 35  
FT CARBOHYD 161 161  
FT CARBOHYD 195 195  
FT CARBOHYD 244 244  
FT CARBOHYD 277 277  
FT CARBOHYD 349 349  
FT VARSPLIC 964 1047  
FT LL -> KADSHSPSLHSOTLETCFSPKGSVGLLA  
FT (IN SHORT ISOFORM).  
FT Q -> E (IN DBSN:5816).  
FT /FPIID-VAR.011968.  
FT T -> S (IN REF.2).  
FT CONFLICT 755 755  
FT SEQUENCE 1047 AA; 117021 MW; 817FB74DEB31F7EF CRC64;  
SO  
Query Match 12.88; Score 460; DB 1; Length 1047;  
Best Local Similarity 41.98; Pred. No.2.6e-24;  
Matches 113; Conservative 39; Mismatches 76; Indels 42; Gaps 6;  
QY 411 DVLIGEQARAD-----GLKRLGK-----LKATLEDA 439  
DB 760 ELVILMERCAADPAERPDGQIKGFIIRNKGSGSIIDNLLRMEQYANNIEKVEER 819  
QY 440 HQALEEKRTKVDLCSTFPCVEAQQIMOGQVQAKKFSNVTLFSDIVGFTAFICQCSF 499  
DB 820 TQAVLEKRAEALVQIQLPHSAEQDKRGFTVQAEAFPSVTIFYFDIYGFTLSMSETP 879  
QY 500 LQVITLMALTYTFDOCCGGLDYKETTIDAYCVAGGL-HKSDTHNOVIALMAIKME 558  
DB 880 MOVVTLINDLYTCDIIDNFDDYKYKETTIDAYVAVSGRNGORHAPETIAMALATL- 938  
QY 559 LSEVMS-----PHGEPIKRIKGLSGSVFAGVGVKMPRYCLFGNNVTLANKPSCSV 612  
DB 939 --DAVSSFRIRNRP-DQLRIRIGVHTGPACAVYGVGKMPRYCLFGDYVTASRMASNGQ 995  
QY 613 PRKINVSPTTYRLILKDCPGFVFPPRSREEL 642  
DB 996 ALKIHVSSTKDALDELGCFQLRGRGVEM 1025

RESULT 14  
ANP\_RAT STANDARD; PRT: 1047 AA.

AC P16067;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Atrial natriuretic peptide receptor B precursor (ANP-B) (ANPRB) (GC-B)  
DE (Guanylate cyclase) (EC 4.6.1.2) (NPR-B) (Atrial natriuretic peptide B-type receptor).  
GN NPR2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=69376566; PubMed=2570641;  
RA Schulz S., Singh S., Bellet R.A., Singh G., Tubb D.J., Chin H., Garbers D.L.;  
RT "The primary structure of a plasma membrane guanylate cyclase demonstrates diversity within this new receptor family.";  
RL Cell 58:1155-1162(1989).  
CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE CYCLASE ACTIVITY ON BINDING OF ANP. SEEMS TO BE STIMULATED MORE EFFECTIVELY BY BRAIN NATRIURETIC PEPTIDE (BNP) THAN BY ANP.  
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C) WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.  
CC -1- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC DOMAIN OF PROTEIN KINASES.  
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; M26896; AAA41205.1; -  
DR PIR; A33300; OYRTBR.  
DR HSSP; Q02846; 1AML.  
DR InterPro; IPR001828; ANP\_receptor.  
DR InterPro; IPR000719; Euk\_Pkinase.  
DR InterPro; IPR001054; G\_cyclase.  
DR InterPro; IPR001170; Ntpep\_receptor.  
DR Pfam; PF00069; Pkinase.1.  
DR Pfam; PF00211; guanylate\_cyc.1.  
DR Pfam; PF01094; ANP\_receptor.1.  
DR PRINTS; PR00255; NATPEPTIDER.  
DR ProDom; PD000001; Euk\_Pkinase.1.  
DR SMART; SM00044; CYC.1.  
DR PROSITE; PS00458; ANP\_RECEPTORS.1.  
DR PROSITE; PS00452; GUANYLATE\_CYCLASES.1.1.  
DR PROSITE; PS50125; GUANYLATE\_CYCLASES.2.1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM.1.  
KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase; cGMP synthesis; Signal.  
FT SIGNAL 1 22  
FT CHAIN 23 1047  
FT DOMAIN 23 458  
FT TRANSMEM 459 478  
FT DOMAIN 479 1047  
FT DOMAIN 513 786  
FT DOMAIN 861 991  
FT DISULFID 75 101  
BY SIMILARITY.

FT DISULFID 439 439 INTERCHAIN (PROBABLE).  
FT DISULFID 448 448 INTERCHAIN (PROBABLE).  
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 349 349 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1047 AA; 117126 MW; 5062C49228C14A3 CRC64;  
Query Match 12.8%; Score 460; DB 1; Length 1047;  
Best Local Similarity 41.9%; Pred. No. 2; be-24;  
Matches 113; Conservative 39; Mismatches 76; Indels 42; Gaps 6;  
OY 411 DVLVIGEOARAD-----GLKRRGK-----LKATPEOA 439  
DB 760 ELVLMERCWADPTERPDRGQIGFRRRKREGSITLNLMLRMQYANLEKVEER 819  
OY 440 HQALEEKRTVDLLCSIFCEVAQQLMGQVQAKFSVNTLMFSDIVGFTALCSQSP 499  
DB 820 TVAYLEERKRAEALVYILPHSVAEQLKRGSTVQAEAFDSVTIVFSIDVFTALASSTP 879  
OY 500 LOVTMINATYTRPDQCGGELDYKVKETIDAYCVAGL-HKESDTAVQIALMALQME 558  
DB 880 MQVTLNDLTYCFDAIINDFDYKVTIGDAIVVSGLRNGQRHAPETARALAL- 938  
OY 559 LSEVMS-----PHGEPIKRIGLHSGSVAGVGVKMPRYCIGNNVTLANFESCSV 612  
DB 939 --DAVSSFRIRHRH-DQLRRIRGVHGPACAGVGLKMPRYCLFGFTVTASMEBSNG 995  
OY 613 PKATVSEPTTYRLKDCPGVFYTPRSREL 642  
DB 996 ALKIVSSTYTKDALDELGCFLQLRGDVEN 1025  
RESULT 15  
ANP\_RAT STANDARD; PRT: 1057 AA.

AC P18910;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)  
DE (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide A-type receptor).  
GN NPR1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Brain;  
RX MEDLINE=69143770; PubMed=2563900;  
RA Chinkers M., Garbers D.L., Chang M.S., Lowe D.G., Chin H., Goeddel D.V., Schulz S.;  
RT "A membrane form of guanylate cyclase is an atrial natriuretic peptide receptor.";  
RT Nature 338:78-83(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91056089; PubMed=1978722;  
RA Yamaguchi M., Rutledge J.J., Garbers D.L.;  
RT "The primary structure of the rat guanylyl cyclase A/atrial natriuretic peptide receptor gene.";  
RT J Biol. Chem. 265:20414-20420(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91352095; PubMed=1679239;  
RA Duda T., Gorczniak R.M., Sharma R.K.;  
RT "Site-directed mutational analysis of a membrane guanylate cyclase cDNA reveals the atrial natriuretic factor signaling site.";





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 12:55:07 ; Search time 9.93048 Seconds

(without alignments)  
2585.358 Million cell updates/sec

Title: US-09-762-767a-4

Perfect score: 3231  
Sequence: 1 MGVNHALELVIRNGPE.....QVFLSRKNTGTEETKODD 619

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3231	100.0	619	1	CYGL_HUMAN
2	3206	99.2	619	1	CYGL_BOVIN
3	3188	98.7	619	1	CYGL_RAT
4	987	30.5	682	1	CYGL_RAT
5	927	28.7	617	1	CYGL_HUMAN
6	916	28.4	730	1	CYGL_RAT
7	909.5	28.1	732	1	CYGL_HUMAN
8	793	24.5	690	1	CYGL_RAT
9	789.5	24.4	690	1	CYGL_HUMAN
10	770	22.8	691	1	CYGL_BOVIN
11	712	22.0	683	1	CYGL_DROME
12	481.5	14.9	1061	1	ANPA_HUMAN
13	468	14.5	1057	1	ANPA_RAT
14	466	14.4	1057	1	ANPA_MOUSE
15	454.5	14.0	1047	1	ANPB_HUMAN
16	453.5	14.0	1047	1	ANPB_BOVIN
17	450.5	13.9	1047	1	CYCD_BOVIN
18	438.5	13.6	1110	1	CYGL_MOUSE
19	433.5	13.4	1108	1	CYGL_MOUSE
20	433	13.4	1109	1	CYGL_MOUSE
21	428.5	13.3	1108	1	CYGL_RAT
22	425	13.2	1050	1	ANPB_ANGUA
23	424	13.1	1108	1	CYGL_HUMAN
24	422.5	13.1	1103	1	CYGL_HUMAN
25	420	13.0	1108	1	CYGL_RAT
26	419	13.0	1103	1	CYGL_BOVIN
27	415.5	12.9	1110	1	CYGL_RAT
28	410	12.7	1137	1	CYGL_CAEEL
29	409	12.6	433	1	KSCG_RAT
30	405.5	12.6	1073	1	HSCR_HUMAN
31	405	12.5	1073	1	HSCR_PIG
32	404	12.5	1072	1	HSCR_RAT
33	391	12.1	1125	1	CYGS_STRPT

34	389	12.0	1076	1	HSCR_CAVPO
35	345.5	10.7	443	1	CYAL_MYCTU
36	309.5	9.6	1134	1	CYAL_BOVIN
37	308.5	9.5	1248	1	CYAL_RAT
38	302.5	9.4	1251	1	CYAL_RAT
39	302.5	9.4	1305	1	CYAL_XENLA
40	301.5	9.3	1249	1	CYAL_HUMAN
41	301.5	9.3	1249	1	CYAL_MOUSE
42	300.5	9.3	2248	1	CYAL_DROME
43	299	9.3	858	1	CYAG_DICDI
44	293	9.1	1144	1	CYAL_HUMAN
45	292.5	9.1	1334	1	CYAL_CHICK

## ALIGNMENTS

RESULT 1  
ID CYGL\_HUMAN STANDARD: PRT; 619 AA.  
AC 002153;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (CCS-beta-1)  
GN GUCY1B1 OR GUCY1B3 OR GUC1B3 OR GUCS83.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Brain;  
RX MEDLINE=92316204; PubMed=1352257;  
RA Guilli G., Scholl U., Bulle F., Gueilleen G.;  
RT Molecular cloning of the cDNAs coding for the two subunits of  
RL Soluble guanylyl cyclase from human brain.";  
FEBS Lett. 304:83-88(1992).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM HSCG-2).  
RC TISSUE-Kidney;  
RX Gansemaans Y., Brouckaert P., Fiers W.;  
RT Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 337-545, AND ALTERNATIVE SPLICING.  
RC TISSUE-Lung;  
RX MEDLINE=92008652; PubMed=1680753;  
RA Chhajlani V., Friendberg P.-A., Ahlner J., Axelsson K.L.,  
RT "Heterogeneity in human soluble guanylate cyclase due to alternative  
RL splicing.";  
FEBS Lett. 290:157-158(1991).  
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.  
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF  
CC MAGNESIUM OR MANGANESE IONS.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HSCG-1 (SHOWN HERE) AND HSCG-2;  
CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE  
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.  
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE  
CC FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X66533; CAA47144.1; -  
 DR EMBL: AF020340; AAB94877.1; -  
 DR PIR: S23097; S23097.  
 DR HSSP: P16068; IAWN.  
 DR Genew: HGNC:4687; GDCY1B3.  
 DR MIM: 139397; -  
 DR InterPro: IPR001054; G\_cyclase.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART: SM00044; CYC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS0125; GUANYLATE\_CYCLASES\_2; 1.  
 DR Lysase: CGMP synthesis; Alternative splicing.  
 KM DOMAIN 421 554 GUANYLATE CYCLASE.  
 FT VARSPLIC 393 425 MISSING (IN ISOFORM HSGC-2).  
 SQ SEQUENCE 619 AA; 70514 MW; 231E4E60DE02A1 CRC64;

Query Match 100.0%; Score 3231; DB 1; Length 619;  
 Best Local Similarity 100.0%; Pred. No. 2e-215;  
 Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYGFVNAHELVIRNYGPEWEDIKKEAQLDEGQFLVRIYDSTKYDLVAASKVYN 60  
 DB 1 MYGFVNAHELVIRNYGPEWEDIKKEAQLDEGQFLVRIYDSTKYDLVAASKVYN 60  
 QY 61 LNAEGLIOMFGKMFVFCQESGYTLIRVAGSNREFLONLDALHDHLATTPGMARPSF 120  
 DB 61 LNAEGLIOMFGKMFVFCQESGYTLIRVAGSNREFLONLDALHDHLATTPGMARPSF 120  
 QY 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAQIHGTEIDMKYIQORNECDHTOFL 180  
 DB 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAQIHGTEIDMKYIQORNECDHTOFL 180  
 QY 122 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAQIHGTEIDMKYIQORNECDHTOFL 180  
 DB 122 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAQIHGTEIDMKYIQORNECDHTOFL 180  
 QY 181 IEKESKEDEYEDLDREENGTOESRISPYTFCKAPPHIIFDRDLVYVQCNAIYRYL 240  
 DB 181 IEKESKEDEYEDLDREENGTOESRISPYTFCKAPPHIIFDRDLVYVQCNAIYRYL 240  
 QY 181 IEKESKEDEYEDLDREENGTOESRISPYTFCKAPPHIIFDRDLVYVQCNAIYRYL 240  
 DB 181 IEKESKEDEYEDLDREENGTOESRISPYTFCKAPPHIIFDRDLVYVQCNAIYRYL 240  
 QY 241 POLQPCNSLISVFSLYRPHIDISFHGISHTNTVYVLSKSGLDVEKLECEDELGTGE 300  
 DB 241 POLQPCNSLISVFSLYRPHIDISFHGISHTNTVYVLSKSGLDVEKLECEDELGTGE 300  
 QY 301 ISCLRKGMYLTPADSILFLCSPSVNMLDLTRKGLYSDIPLHDATRDVLVGEOPR 360  
 DB 301 ISCLRKGMYLTPADSILFLCSPSVNMLDLTRKGLYSDIPLHDATRDVLVGEOPR 360  
 QY 361 EEKKTQELIETLDRLQLTLRALEDEKKTDTLLSVLPVYANELRKRPPAKRYDNV 420  
 DB 361 EEKKTQELIETLDRLQLTLRALEDEKKTDTLLSVLPVYANELRKRPPAKRYDNV 420  
 QY 421 TILFSGIVGFNAFCSHASGEGAMKIVNLDLYRFDLTDSRKRPFYKYEYGDYK 480  
 DB 421 TILFSGIVGFNAFCSHASGEGAMKIVNLDLYRFDLTDSRKRPFYKYEYGDYK 480  
 QY 481 TVSGIPEPCIHARSHICHALDMMETAGOVQVDSGVOITIGIHGEVNTGVYIGORMPY 540  
 DB 481 TVSGIPEPCIHARSHICHALDMMETAGOVQVDSGVOITIGIHGEVNTGVYIGORMPY 540  
 QY 541 CLFGNTVNTSRFTTGEKGINVSEYTYRCLMSPNSDOPFHLERGPVSKGKKEPMQ 600  
 DB 541 CLFGNTVNTSRFTTGEKGINVSEYTYRCLMSPNSDOPFHLERGPVSKGKKEPMQ 600  
 QY 601 VWFELSKNTGTEETKODDD 619  
 DB 601 VWFELSKNTGTEETKODDD 619

RESULT 2  
 CYG1\_BOVIN STANDARD; PRT; 619 AA.  
 AC P16068;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (GCS-beta-1)

DE (Soluble guanylate cyclase small subunit).  
 GN GUCY1B1 OR GUCY1B3 OR GUCY1B3.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=lung;  
 RX MEDLINE=89031214; PubMed=2903071;  
 RA Koesling D., Herz J., Gausepohl H., Niroomand F., Hirsch K.-D.,  
 Muelsch A., Boelme E., Schultz G., Frank R.;  
 RT "The primary structure of the 70 kDa subunit of bovine soluble  
 guanylate cyclase."  
 RL FEBS Lett. 239:29-34 (1988).  
 RN [2]  
 RP 3D-STRUCTURE MODELING OF 412-572.  
 RX MEDLINE=96054247; PubMed=9391039;  
 RA Liu Y., Ruoho A.E., Rao V.D., Hurley J.H.;  
 RT "Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling  
 and mutational analysis."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419 (1997).  
 CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.  
 CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF  
 MAGNESIUM OR MANGANESE IONS.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: LUNG AND BRAIN.  
 CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE  
 FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.  
 CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE  
 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: Y00770; CAA68739.1; -  
 DR PIR: S01653; OYB070.  
 DR PDB: 501653; OYB070.  
 DR InterPro: IPR001054; G\_cyclase.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART: SM00044; CYC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS0125; GUANYLATE\_CYCLASES\_2; 1.  
 DR Lysase: CGMP synthesis; 3D-structure.  
 KM DOMAIN 421 554 GUANYLATE CYCLASE.  
 FT SEQUENCE 619 AA; 70502 MW; 8EFB14952B80F344 CRC64;

Query Match 99.2%; Score 3206; DB 1; Length 619;  
 Best Local Similarity 99.0%; Pred. No. 1e-213;  
 Matches 613; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 MYGFVNAHELVIRNYGPEWEDIKKEAQLDEGQFLVRIYDSTKYDLVAASKVYN 60  
 DB 1 MYGFVNAHELVIRNYGPEWEDIKKEAQLDEGQFLVRIYDSTKYDLVAASKVYN 60  
 QY 61 LNAEGLIOMFGKMFVFCQESGYTLIRVAGSNREFLONLDALHDHLATTPGMARPSF 120  
 DB 61 LNAEGLIOMFGKMFVFCQESGYTLIRVAGSNREFLONLDALHDHLATTPGMARPSF 120  
 QY 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAQIHGTEIDMKYIQORNECDHTOFL 180  
 DB 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAQIHGTEIDMKYIQORNECDHTOFL 180  
 QY 122 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAQIHGTEIDMKYIQORNECDHTOFL 180  
 DB 122 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAQIHGTEIDMKYIQORNECDHTOFL 180  
 QY 181 IEKESKEDEYEDLDREENGTOESRISPYTFCKAPPHIIFDRDLVYVQCNAIYRYL 240  
 DB 181 IEKESKEDEYEDLDREENGTOESRISPYTFCKAPPHIIFDRDLVYVQCNAIYRYL 240



```

OY 241 POLDPGNCSSLVSFLVRPHIDISFHGILSHINTVFLRSKGLLDVKECEDELGTGE 300
DB 241 POLDPGNCSSLVSFLVRPHIDISFHGILSHINTVFLRSKGLLDVKECEDELGTGE 300
OY 301 ISCLRLKGMITYLPREADSILFLCSPSVNMDDLTRRGYLSIPPLHADRDLVLLGEOR 360
DB 301 ISCLRLKGMITYLPREADSILFLCSPSVNMDDLTRRGYLSIPPLHADRDLVLLGEOR 360
OY 361 EBYKLTQELIELDRLQTLRALDEKKKTDLTLSVPPSVANELRHRKRPVAKRYDNV 420
DB 361 EBYKLTQELIELDRLQTLRALDEKKKTDLTLSVPPSVANELRHRKRPVAKRYDNV 420
OY 421 TILFSGIVGFNAFCSKHAASGEGAMKIVNLNDLYTRFDTLDSRKNPFYKVEYGDKYM 480
DB 421 TILFSGIVGFNAFCSKHAASGEGAMKIVNLNDLYTRFDTLDSRKNPFYKVEYGDKYM 480
OY 481 TVSGLPFPCIHARSICHALDMMELIAGOVYDGSVOITGIRHGEVVTGYIGORMPRY 540
DB 481 TVSGLPFPCIHARSICHALDMMELIAGOVYDGSVOITGIRHGEVVTGYIGORMPRY 540
OY 541 CLFGNTVNLSTRTETTGKGINVSEYTYRCIMSPENDPOPHLEHRGPVSKGKKEPMQ 600
DB 541 CLFGNTVNLSTRTETTGKGINVSEYTYRCIMSPENDPOPHLEHRGPVSKGKKEPMQ 600
OY 601 VWFLSRKNTGTEETKODDD 619
DB 601 VWFLSRKNTGTEETKODDD 619

RESULT 3
CYG1_RAT STANDARD: PRT: 619 AA.
AC P20595;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanylate cyclase soluble, beta-1 chain (EC 4.6.1.2) (GCS-beta-1)
GN GUCY1B1 OR GUCY1B3 OR GUC1B3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89087429; PubMed=2905128;
RA Nakane M., Sahel S., Kuno T., Ishii K., Murad F.;
RT "Molecular cloning of a cDNA coding for 70 kilodalton subunit of
RT soluble guanylate cyclase from rat lung."
RL Biochem. Biophys. Res. Commun. 157:1139-1147(1988).
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: LONG AND BRAIN.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL, M22562; AAA41204.1;
CC PIR, A31871; OYFBL.
CC HSP, P16068; IAWN.

```

```

DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYC; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS00125; GUANYLATE_CYCLASES_2; 1.
KW Lyase; GMP synthesis.
FT DOMAIN 421 554
SQ SEQUENCE 619 AA; 70477 MW; BCC9850DE4504226 CRC64;

Query Match      98.7%; Score 3188; DB 1; Length 619;
Best Local Similarity 98.5%; Pred. No. 1,8e-212;
Matches 610; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 MTGFVNHAEILVIINRYGPEWEDIKKEAQDDEGOFRIITYDSKTYDLVAASKVYN 60
DB 1 MTGFVNHAEILVIINRYGPEWEDIKKEAQDDEGOFRIITYDSKTYDLVAASKVYN 60
OY 61 INAGELIOMFGMFVFCOESGYDTILRVLSNVREFTQNLDAHDHLATIPGARAPSF 120
DB 61 INAGELIOMFGMFVFCOESGYDTILRVLSNVREFTQNLDAHDHLATIPGARAPSF 120
OY 121 RCTDAEKGGLIHYSEREGADIVIGIKTVAQOIHGETIDMKVYIQORNECHTOPL 180
DB 121 RCTDAEKGGLIHYSEREGADIVIGIKTVAQOIHGETIDMKVYIQORNECHTOPL 180
OY 181 IEKESKEDEFEYEDLDREENGTOESRISPYTFCKAFPHIIFDRDLVYTCGNATYRVL 240
DB 181 IEKESKEDEFEYEDLDREENGTOESRISPYTFCKAFPHIIFDRDLVYTCGNATYRVL 240
OY 241 POLDPGNCSSLVSFLVRPHIDISFHGILSHINTVFLRSKGLLDVKECEDELGTGE 300
DB 241 POLDPGNCSSLVSFLVRPHIDISFHGILSHINTVFLRSKGLLDVKECEDELGTGE 300
OY 301 ISCLRLKGMITYLPREADSILFLCSPSVNMDDLTRRGYLSIPPLHADRDLVLLGEOR 360
DB 301 ISCLRLKGMITYLPREADSILFLCSPSVNMDDLTRRGYLSIPPLHADRDLVLLGEOR 360
OY 361 EBYKLTQELIELDRLQTLRALDEKKKTDLTLSVPPSVANELRHRKRPVAKRYDNV 420
DB 361 EBYKLTQELIELDRLQTLRALDEKKKTDLTLSVPPSVANELRHRKRPVAKRYDNV 420
OY 421 TILFSGIVGFNAFCSKHAASGEGAMKIVNLNDLYTRFDTLDSRKNPFYKVEYGDKYM 480
DB 421 TILFSGIVGFNAFCSKHAASGEGAMKIVNLNDLYTRFDTLDSRKNPFYKVEYGDKYM 480
OY 481 TVSGLPFPCIHARSICHALDMMELIAGOVYDGSVOITGIRHGEVVTGYIGORMPRY 540
DB 481 TVSGLPFPCIHARSICHALDMMELIAGOVYDGSVOITGIRHGEVVTGYIGORMPRY 540
OY 541 CLFGNTVNLSTRTETTGKGINVSEYTYRCIMSPENDPOPHLEHRGPVSKGKKEPMQ 600
DB 541 CLFGNTVNLSTRTETTGKGINVSEYTYRCIMSPENDPOPHLEHRGPVSKGKKEPMQ 600
OY 601 VWFLSRKNTGTEETKODDD 619
DB 601 VWFLSRKNTGTEETKODDD 619

RESULT 4
CYG2_RAT STANDARD: PRT: 682 AA.
AC P22717;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanylate cyclase soluble, beta-2 chain (EC 4.6.1.2) (GCS-beta-2).
GN GUCY1B2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE-KIDNEY;
RX MEDLINE-9105012; PubMed-1980215;
RA Yuen P.S.T., Potter L.R., Garbers D.L.;
RT "A new form of guanylyl cyclase is preferentially expressed in rat kidney."
RL Biochemistry 29:10872-10878(1990).
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF MAGNESIUM OR MANGANESE IONS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: KIDNEY AND LIVER.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M57507; AAA41207.1; -
DR PIR; A36228; OYRTB2.
DR HSSP; P19687; 1AMN.
DR InterPro; IPR001054; G_cyclase.
DR InterPro; IPR001230; Prenyl_site.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR Lyase; CGMP synthesis.
KW DOMAIN 408 536 GUANYLATE CYCLASE.
SQ SEQUENCE 682 AA; 76196 MW; 98C173C1A1CC7715 CRC64;
Query Match 30.5%; Score 987; DB 1; Length 682;
Best Local Similarity 36.9%; Pred. No. 1,4e-60;
Matches 225; Conservative 104; Mismatches 195; Indels 86; Gaps 12;
QY 66 ILDMGKMFVPEQSGDYTLIRVGSNVEFLQNTDALDHLATIPGRAPSFRTDA 125
DB 4 ILDFGEYFKFKMSGYDMLKRLTGLNTEFLNLDALSYALSTQENMABSFREEG 63
QY 126 EKSGGLIATYSEREGADIVIGIKTYAAQIHGTEDMKVIOQRNEC-----DHTQF 179
DB 64 ADG-AMLIHYSDRHGCHIVPGIIEAVAKDFPTDVAMSLDM-NEVERTGKEHYVF 121
QY 180 LIEBKESKE-----EDFYED----- 194
DB 122 LVVQAKARQIRGAKASRQSGEDSQADQALQGLTLKMKERYINIPVCGEKSHSTAVRA 181
QY 195 -----LDREENTQGSRSISPYFFCAFPHTIEDVDVYTOGGMAYRPLQDP 245
DB 182 SVLFGGLDQTPYVPERLWEEVEYCAFPHYFDALRKQAGVNIQKVPILTF 241
QY 246 GNCGLSVGLVPHIDISFHLISHTVFLVSKGLDVEKCEDELTGETEISCLR 305
DB 242 QKRALDEYFIIHPVTFNLSICKFINSQFVLTREMPKAR-----KSGPMK 292
QY 306 LKQOMIYLPKADSIPLCSPSVANLDDLTRRGLYSLDIPLDATRDILVLLGEQFREERYKL 365
DB 293 LRQOMIMESLRQIMFICSPSVNSIQLIEBSKMLSDIADPHDTTRDILINQOLAMEL 352
QY 366 TOLELEITDRLQTLRLADEKKRTDILKSVLPSPVANLRRKRPAPARYDNVTLFS 425
DB 353 SCLEKKKEELRYLSNHLATEKKRTETLLYAMLPEHYANOLKGRKVAAGEFETCTILFS 412
QY 426 GIYGFNAFCSHASGEMAKIVNLNDLYTRPDLTFDSRKNPFFYKVEYEGDKYMTVSGL 485
DB 413 DVTFTFNIC---AAFC-PIQIVYMLNSMYSKFRLTSVHD---YKVEYTGDAIMVYGVV 465

```

```

QY 486 PEPCHHRSICHLALDMEIAGVQ--VDGESVQITIGITGEVTVGVIGOMPRYCLF 543
DB 466 PVPESHAGVAVNPAFALGMRISAKENVMPVGEPIQIVGHTSPVLAVGVDGDMPRCLF 525
QY 544 GNYVNLTSRTETTGEEKKIVSEYTRCLMSPENDPQFLHHRGYPYSGKREPMQWVF 603
DB 526 GDTYVNLASRRESQGLPSKVLSPVAHRAL-----KMGFEIVRGEIEVNGKGR-MTTF 579
QY 604 LSRKNTGEE 613
DB 580 LIQNLNATED 589
RESULT 5
CYG2 HUMAN
ID CYG2 HUMAN STANDARD; PRT; 617 AA.
AC 075343; Q9NZ64;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanylate cyclase soluble, beta-2 chain (BC 4.6.1.2) (GCS-beta-2).
GN GUCY1B2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A., AND REVISION TO 175.
RC TISSUE-Kidney;
RX MEDLINE-20241821; PubMed-10777682;
RA Behrends S., Vense K.;
RT "The beta(2) subunit of soluble guanylyl cyclase contains a human-specific frameshift and is expressed in gastric carcinoma."
RL Biochem. Biophys. Res. Commun. 271:64-69(2000).
RN 121
RP SEQUENCE OF 95-257 FROM N.A.
RC TISSUE-Heart;
RX MEDLINE-99107820; PubMed-9889008;
RA Behrends S., Kazmierczak B., Steenpa A., Knauf B., Bullerdiek J., Scholz H., Elberg H.;
RT "Assignment of GUCY1B2, the gene coding for the beta2 subunit of human guanylyl cyclase to chromosomal band 13q14.3 between markers D13S168 and D13S155."
RL Genomics 55:126-127(1999).
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF MAGNESIUM OR MANGANESE IONS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF038499; AAD09440.2; -
DR EMBL; AF218383; AAF66105.1; -
DR HSSP; P19687; 1AMN.
DR Genew; HGNC:4686; GUCY1B2.
DR KIM; 603695; -
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; FALSE_NEG.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.

```



RESULT 7			
ID	CYGA_HUMAN	STANDARD:	PRT: 732 AA.
AC	P33402.		
DT	01-FEB-1994 (Rel. 28, Created)		
DT	01-FEB-1994 (Rel. 28, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Guanylate cyclase soluble, alpha-2 chain (EC 4.6.1.2) (GCS-alpha-2).		
CN	GUCYLA2 OR GUCL1A2 OR GUCSA2.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=92070494; PubMed=1683630;		
RA	Hartenack C., Wedel B., Koesling D., Malkevitz J., Boehme E.,		
RT	Schultz G.;		
RT	"Molecular cloning and expression of a new alpha-subunit of soluble		
RT	guanylyl cyclase. Interchangeability of the alpha-subunits of the		
RT	enzyme.";		
RL	FEBS Lett. 292:217-222(1991).		
RL	[2]		
RN	SEQUENCE FROM N.A. (ISOFORM ALPHA 2-1).		
RP	MEDLINE=95403342; PubMed=7673142;		
RX	Behrends S., Hartenack C., Schultz G., Koesling D.;		
RA	"A variant of the alpha 2 subunit of soluble guanylyl cyclase		
RT	contains an insert homologous to a region within adenylyl cyclases		
RT	and functions as a dominant negative protein.";		
RT	J. Biol. Chem. 270:21109-21113(1995).		
CC	-1- FUNCTION: HAS GUANYLYL CYCLASE ON BINDING TO THE BETA-1 SUBUNIT.		
CC	THE ALTERNATIVE SPLICED ISOFORM, ALPHA-2-I, ACTS AS A NEGATIVE		
CC	REGULATOR OF GUANYLYL CYCLASE ACTIVITY AS IT FORMS NON-FUNCTIONAL		
CC	HETERODIMERS WITH THE BETA SUBUNTS.		
CC	-1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.		
CC	-1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF		
CC	MANGNESIUM OR MANGANESE IONS.		
CC	-1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.		
CC	-1- SUBCELLULAR LOCATION: CYTOSOLSMC.		
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA-2 (SHOWN HERE) AND		
CC	ALPHA-2-I; ARE PRODUCED BY ALTERNATIVE SPLICING.		
CC	-1- TISSUE SPECIFICITY: ALPHA-2 ISOFORM IS EXPRESSED IN FETAL BRAIN,		
CC	LIVER, COLON, ENDOTHELUM AND TESTIS, WHILE THE ALPHA-2-I FORM		
CC	IS EXPRESSED ONLY IN LIVER, COLON AND ENDOTHELUM.		
CC	-1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLYL CYCLASES: SOLUBLE		
CC	FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.		
CC	-1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE		
CC	FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; X63382; CAA444921.1; -		
DR	EMBL; Z50053; CAA90393.1; -		
DR	PIR: S18325; S18325.		
DR	HSSP; P19687; 1AMN.		
DR	GeneW; HGNC:4684; GUCYLA2.		
DR	MIM; 601244; -		
DR	InterPro; IPR001054; G_cyclase.		
DR	Pfam; PF00211; guanylate_cyc; 1.		
DR	SMART; SM00044; CYC; 1.		
DR	PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.		
DR	PROSITE; PS01253; GUANYLATE_CYCLASES_2; 1.		
FW	Lyase; CGMP synthesis; Multigene family; Alternative splicing.		
FT	DOMAIN 51 648		
FT	DOMAIN 51 76		
FT	DOMAIN 51 58		
FT	POLY-ALA.		

FT	FT	VARSP/LIC	612	612	Q -> OPOSSIBLE PVSIDVDPHQSEFIDATEK (IN ISOPFORM ALPHA 2-1)
50	SEQUENCE	732 AA;	81749 MW;	79465A1D3FE5DB7	CNC64;
	Query Match	28.1%;	Score 909.5;	DB 1;	Length 732;
	Best Local Similarity	37.7%;	Pred. No. 3.3e-55;		
	Matches 208;	Conservative 111;	Mismatches 175;	Indels 57;	Gaps 14
QY	51	LVAASVTVLNLANGELLQFMGKMFVPOQESDYDILNLAVLSNREFLQNDALDHILAT	110		
DB	154	ILDQTNAILGKLFKEELQKRFGEEFFNYICHEN-ENVLAIVAGTLLQDPEFNGFDALHEIRT	212		
QY	111	IV---PGRMAPSRCTDAERKAGKGLILHYHSENEGQDQIVICQILITVAQOILNGEIDEIKVI	167		
DB	213	SFGKQATLSPSP/LCKELPEGT-LMLHFVPHHIVGFMQMGIMAKAGKKIY-RLDVEYE	269		
QY	168	QQRNE-----CDITQFLIEKSKEDFEEDLDREENGTOESRISPTPEKAF	217		
DB	270	QVANEKICSDVSNPGNSCTFLI--KECENTNIMKMLPOSTQVPPADLRISINTFCRAF	327		
QY	218	PFHIFDRDLVYQCGNATVRYLPOLQPGNSCLSV-----FSIVRPHIDISPHGILSH	271		
DB	328	PHLMEDPSNIVQLQLEGRKQD-----RCDTHVYLKEDCFELVSPKNATFERILR	381		
QY	272	INTFVPLRSKGLLDVKECEDELGTPE--ISCLRKGMQIYLPREADSILFLCSPSVN	329		
DB	382	LSHPVYIRTK-----PEASGSENNKRVAEVVGQMIHVPESSNLSILFSGPCVDK	429		
QY	330	LDLTFRRGYLSDPLPHDTRDLVILRGDFREYVLTQDELINDRLOLTIRALEDEKK	389		
DB	430	LDLMDGRGLHSDIPRHDTRVYILVGEDAKAQDGLKRMKRLATLERHQALEEKK	489		
QY	390	TDLFLYSLVPPSVANELRRKRPVPAKRDYNTILFSGIVGFNAFCSRKASGEMKIVNL	449		
DB	490	TVDLTSLFPPGDAQQLMGQOQVAKRKEDVYMLFSDIVGTALCACT---PMQVISM	545		
QY	450	LNDLYRPFQLDLSRKNPF--YKRYETVGDKNYVSGLRPECIHAHSICHLADMEIA	507		
DB	546	LNELYTRFD-----HCGFLDIYKVEYTGDAICVAAGLHRRSLCHAPRIALMAKHELS	600		
QY	508	GQVQV-DGESVOITGIIHTEGEVTVGIVGQMPRYCLFGNTVNLSTRTEGKINKIVSE	566		
DB	601	EEVLTDPDRPIQRIRIGIHSGSVLACVAGVBMRYCLFGNNVTLASKESSGHPRIIVSP	660		
QY	567	YTYRCMLSPEN 577			
DB	661	TYQLLKRRES 671			
RESULT 8					
CY3G_RAT	ID	CY3G_RAT	STANDARD;	PRT;	690 AA.
AC	P19686;				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Guanylate cyclase soluble, alpha-1 chain (EC 4.6.1.2) (GCS-alpha-1)				
DE	(Soluble guanylate cyclase large subunit).				
GN	GUCYA1L OR GUCYAL.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=101116;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE=Lung;				
RX	MEDLINE=91009100; PubMed=1698769;				
RA	Nakane M., Arai K., Saheki S., Kuno T., Buechler W., Murad F.;				
RT	"Molecular cloning and expression of cDNAs coding for soluble				
RL	guanylate cyclase from rat lung."				
RL	J. Biol. Chem. 265:16841-16845(1990).				
RP	[2]				
RP	SEQUENCE FROM N.A.				

```

RC STRAIN-Sprague-Dawley; TISSUE-Striatum;
RA MEDLINE-97151525; PubMed-8997507;
RX Smigrodzki R.M., Levitt P.;
RT "The alpha 1 subunit of soluble guanylyl cyclase is expressed
RT prenatally in the rat brain."
RL Brain Res. Dev. Brain Res. 97:226-234(1996).
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M57405; AAA41206.1; -
DR EMBL: U60835; AAB17953.1; -
DR PIR: A38297; OYR1A1.
DR HSSP: P19687; 1AMN.
DR InterPro: IPR001054; G_CYCLASE.
DR Pfam: PF00211; guanylate_cyc; 1.
DR SMART: SM00044; CYCG; 1.
DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 1.
DR Lysase; CGMP synthetase; Multigene family.
KM DOMAIN 480 607 GUANYLATE CYCLASE.
FT SEQUENCE 690 AA; 77566 MW; E4819B2CAAF86401 CRC64;
SQ
Query Match 24.5%; Score 793; DB 1; Length 690;
Best Local Similarity 34.7%; Pred. No. 3.3e-47;
Matches 214; Conservative 100; Mismatches 215; Indels 88; Gaps 20;
OY 27 KRAQDESGPLVRIYIDSKTYDYAAASKVLNLAGELIC-MGCKMFVYQESGYDT 85
DB 102 KENRNSSEKEDERITAEBA-----TACGVPV-----EVLDSJGEELFKCYEED-EH 149
OY 86 ILRLVGSNVRREFLQNDAL-----HDHLATIPGMRAPSFRTDAEKGKGLILHYSERE 140
DB 150 ILGVVGGLTKDPLNFSTLLKQSHQGEERGRLEDAIICLDKQ-DFLANYYTFPPRR 208
OY 141 GLQDIVIGITVAQOIHGTEDMKVYIQ-RNEECDHTQFL-----IEEKSEKEDF 191
DB 209 TTAALLPGITKAAARILYSEHVEVSLMPCFSECC--TEFVQPYLLGVHVKSTKPS-- 264
OY 192 YEDLDREFENGQOESRISPYT-FCKAPPHIIFDRDLYVYQGNALYRVLQQL-PCNGS 249
DB 265 -----LSFGKQSSVAVPTSLFCNTPPHFMDRLALLOLGNGIRLVNKRDFGKPN 318
OY 250 LLSVSLVPHIDISFNGHLSHINTVFVLRSKEGLDVEKCEDELTEISCLRLKQ 309
DB 319 PEEFEILLPKINQFSGINTMLNMQFVIRVRW-----DNLVAKSSVMDLQKQ 368
OY 310 MIYLPDADSIPLFCSPSVANNLDDLRRGLYLSIDIPHLATRDVLVLGEGFREFRYKYLQEL 369
DB 369 MIYIYESSALIFLFGPCVDRLDEDFGRGLYLSIDIPHNALRYVLLGEGQARADGLKRL 428
OY 370 ELTDRIGLQTLRALDEKKTDTLLYSVPSPSVANLRRKRPVAPARVYNVITFSGIYG 429
DB 429 GKLKATLTAHQAOLDEKKKIVDLCSTIPSEVAQOLGQVQAKKFEVTLMLFSDIYG 488
OY 430 FNAFOSKSHASGAGAKIVMLNDLYTRFDTLTDSRKNPFYKVEYGDYKMTYSGJPEPC 489
DB 489 FTAICSGCS-----PLQVITMLNALTTRFQGGCELD--VYKVEITIGDAYCAVAGIHRRES 541

```

```

OY 490 IHNASICHALDMMELAGOV-QVDESGVQITIGITGEVTVGICQRMPPRCLEGNVYN 548
DB 542 DTHAVQIALMALMKELSNESVPHGEPIKMRIGLSSGVFAGVGVKMPRCLEGNVYN 601
OY 549 LRSRTTEGKKGKINSEYTRCL-----NSPENSDFQF-----HL-----EH 586
DB 602 LANKFSCSVPRKINSPYTRLLKDCPGFVFTRPSRELPLNPFPSDIPGICHLFDAYOH 661
OY 587 RGPVSMKGRKEPMQVWF 603
DB 662 QGPNs-----KPMF 670
OY
DB
RESULT 9
CYG3_HUMAN
ID CYG3_HUMAN STANDARD; PRT; 690 AA.
AC 002108; 043843;
DT 01-JUL-1993 (Rel. 26, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanylate cyclase soluble, alpha-1 chain (EC 4.6.1.2) (GCS-alpha-1)
DE (Soluble guanylate cyclase large subunit) (GCS-alpha-3).
GN GUCY1A1 OR GUCY1A3 OR GUC1A3 OR GUCS3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-92316204; PubMed-1352257;
RA Guaili G., Scholl U., Buile F., Guellaen G.;
RT Molecular cloning of the cDNAs coding for the two subunits of
RT soluble guanylyl cyclase from human brain..";
RL FEBS Lett. 304:83-88(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA Ganssems Y., Brouckaert P., Fiers W.;
RT "Human soluble guanylate cyclase large subunit mRNA, alpha3-like.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-98416113; PubMed-9742212;
RA Zabel U., Weeger M., La M., Schmidt H.H.;
RT "Human soluble guanylate cyclase: functional expression and revised
RT isoenzyme family.";
RL Biochem. J. 335:51-57(1998).
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- ENZYME REGULATION: ACTIVATED BY NITRIC OXIDE IN THE PRESENCE OF
CC MAGNESIUM OR MANGANESE IONS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF GUANYLATE CYCLASES: SOLUBLE
CC FORMS AND MEMBRANE-ASSOCIATED RECEPTOR FORMS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X66534; CAA47145.1; -
DR EMBL: U58855; AAB94794.1; -
DR EMBL: Y15723; CAA75738.1; -
DR PIR: S23098; S23098.
DR HSSP: P19687; 1AMN.
DR Genew; HGNC:4685; GUCY1A3.

```





```

OY 230 TQCGNAIYRVLPOLQ-PCNCSILSVSLVRPHIDISFGLISHINTVTFVLRSEGLIDYE 288
DB 300 LQLGIRLRMRVQVQKPEDEFEYELIPKISQTFSGIMTLMNQFLVVRWR----- 354
OY 289 KLECEDELGTETISCLRLKGMITYLEADSIIFLCSPSWNTDDLRRLGYISDIPLDA 348
DB 355 -----DNSMKSSRWMDLKGOMITYWESSILFLSPCYDRLEDFGRGLYSDIPIHNA 409
OY 349 TRDLVLLGEQREERYKLTQLEELIDRLQLTLRALDEDEKKDTLLSYLPSVANELRH 408
DB 410 LADVVLIGQARAOGLKRLGLKATLEQAHQALEEERKRYVDLCLSFSPVANQIMQ 469
OY 409 KRPVAKRKDNTIIFSGIVGNACSKASGEGAMKIYNLDLTYRDTLTDKSKNP 468
DB 470 GHAVQAKRGNTMTLFSIDVFTALCSQCS-----PLQVITMLNALTREFDROCGELD--- 522
OY 469 YKKEVTEGDKYMTVSGELPEPCIHARSICHLADMEIAGO-VQVDSGVQITIGHGE 527
DB 523 YKVEITIGDAYCVAGGLHAKESDTHAVQIALMLKMMELSHVYSPGEPKMTIGLHSGS 582
OY 528 VVTGVIGORMPRYCLEGNTVNLSTRTETTGKINVSERYRCL-----NSPEN 577
DB 583 VVAGVGVGMRYCLFENNVTLANKEFSCSVPRKINVSPTYRLKDCGFVFTPRSRBE 642
OY 578 SDPOF 582
DB 643 LPPNF 647

RESULT 11
CYGH_DROME STANDARD; PRT; 683 AA.
ID CYGH_DROME
AC 007093;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Head-specific guanylate cyclase (EC 4.6.1.2).
GN GYC-ALPHA-99B OR GYC-ALPHA-63A OR GYC OR DGCL.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Braachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canlon-S; TISSUE=Head;
RX MEDLINE=93203896; PubMed=8095978;
RA Yoshikawa S., Miyamoto I., Aruga J., Furutachi T., Okano H.,
RA Mikoshiba K.,
RT "Isolation of a Drosophila gene encoding a head-specific guanylyl
RT cyclase."
RL J. Neurochem. 60:1570-1573(1993).
CC -1- FUNCTION: MAY HAVE A ROLE IN PHOTOTRANSDUCTION. A SECOND SUBUNIT
CC MAY BE REQUIRED FOR ENZYME ACTIVITY.
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- SUBUNIT: DIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- TISSUE SPECIFICITY: HEAD, WHERE IT IS PREFERENTIALLY EXPRESSED IN
CC THE CNS AND THE RETINA. NOT FOUND IN BODIES.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S57126; AAB35820.1; -
CC DR PIR: JH0810; JH0810.
CC HSP: P19687; IAWW.

```

```

DR Flybase: FBgn0013972; Gyc-alpha-99B.
DR InterPro: IPR001054; G_cyclase.
DR Pfam: PF00211; guanylate_cyc_1.
DR SMART: SM00044; CYCC_1.
DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; FALSE_NEG.
DR PROSITE: PS01125; GUANYLATE_CYCLASES_2; 1.
KW Lyase; GMP synthetase; Multigene family; Vision.
FT DOMAIN 465 591 GUANYLATE CYCLASE.
SQ SEQUENCE 683 AA; 75906 MW; 12D67CAB79A07C8D CRC64;

Query Match 22.0%; Score 712; DB 1; Length 683;
Best Local Similarity 30.2%; Pred. No. 1,3e+41;
Matches 188; Conservative 121; Mismatches 229; Indels 84; Gaps 20;

OY 22 MEDIKREADLDEEGQLVRIIYDSDTYDLVAASKYLV-----NLNAGEILOMGKKEFEV 77
DB 73 WPIRIH-LKLDPO-----TFKSCANYDIADIDQELLMKMDASSEIIVLGEELITC 124
OY 78 COESGYDTLLRVLSGVNREFLQMLDHLHDLATVYGMRAVPSFCTDAEKGLILHYYS 137
DB 125 CCTGITERAFRCIGTDLQELFGLSDGVYDLKLQEDVDYDTGVCA-----GEGELL-FTS 179
OY 138 EREGLODIYIGIKYTAQOIHGEIEMKVYQGNNEC-----DHTQFLIE 182
DB 180 ERPVIALMLLGSLSKALTFRMLYKVDVNIKI--EVEKCPAVSLPLAGQGQLPHAD--G 234
OY 183 EKSEKDEFEYEDDREENGTOESRISPYTFCAPFPHITFDRLVYTOGNAIYRVLP- 241
DB 235 SSTSVSKTIPETVQGRNSSNASDLQNNSSFCGMFWHFMNQLEVLGRFSILYKP 294
OY 242 QLOPGNCSLSVSLVRPH-IDISFGLISHINTVTV-----LRSKGLLDVKEKCEDELNG 298
DB 295 YMADEFCQATVYDFEFRPKGLTKFRFDIYRRTYPLFGLINPNPQAVDFPAIG----- 347
OY 299 TELSCIRLKGOMIYLEADSIIFLCSPSWNTDDLRRLGYISDIPLHATRLVLLGEQ 358
DB 348 -----LEIKGOMVHCPESNSLFTSGPPLDGLGTCNGELISDIPLDHATREVLIVGEQ 402
OY 359 FREERYKLTQLEELIDRLQLTLRALDEDEKKDTLLSYVLPSPVANELRHRRKRPVAKRYD 418
DB 403 ARQDGLRRMRDMKIKNSIEEANSAYKERRKNVSLHLIFPAIARKNMVGSSIDAKTYP 462
OY 419 NVTILFSGIVGNAPCSKASGEGAMKIYNLDLTYRDTLTDKSKNP--VYKVEYVG 476
DB 463 DVTILFSDIYGFYSICSRAT---PPMVISMELGLEYDEFEDEF---FFDYKVEITIG 513
OY 477 DKMYTVSGELPEPCIHARSICHLADMEI-----AQGVYDGSVQITIGHGEYVTVG 532
DB 514 DATCVASGLHRASITDA---HRCLDGLKMDACSKHITHDGIKRMKRGILHGTVLAVG 569
OY 533 IGRMPRYCLFGNTVNLSTRTETTGKINVSERYRCLMSPENDPOPHLHRGPVSM 592
DB 570 VGRKMRRYCLFGISVITANKFESGSEALKINVSPTTKDWLTKEGF--PELOPRDP-SF 626
OY 593 KGRKPMQVWFSLRKNTGTEET 614
DB 627 LPKEFP-----NPGGTEY 639

RESULT 12
ANPA_HUMAN STANDARD; PRT; 1061 AA.
ID ANPA_HUMAN
AC P16066;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)
DE (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide
DE A-type receptor).
GN NPRI OR ANPRA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=89356605; PubMed=2569967;  
 RA Lowe D.G., Chang M.S., Hellmias R., Chen E., Singh S., Garbers D.L.,  
 RA Goeddel D.V.;  
 RT "Human atrial natriuretic peptide receptor defines a new paradigm for  
 RT second messenger signal transduction.";  
 RL EMBO J. 8:1377-1384(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=98289600; PubMed=9618281;  
 RA Takahashi Y., Nakayama T., Soma M., Izumi Y., Kamatsuse K.;  
 RT "Organization of the human natriuretic peptide receptor A gene";  
 RL Biochem. Biophys. Res. Commun. 246:736-739(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Maeda N., Knowles J.W.;  
 RT "Identification of functional polymorphisms in noncoding regions of  
 RT the human natriuretic peptide receptor A gene";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 634-1048 FROM N.A.  
 RC TISSUE=Retina;  
 RX MEDLINE=95042574; PubMed=7954658;  
 RA Pardhasaradhi K., Kutty R.K., Gentileman S., Krishna G.;  
 RT "Expression of mRNA for atrial natriuretic peptide receptor guanylate  
 RT cyclase (ANPRA) in human retina";  
 RL Cell. Mol. Neurobiol. 14:17(1994).  
 CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE  
 CC CYCLASE ACTIVITY ON BINDING OF ANP.  
 CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO  
 CC WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)  
 CC WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE  
 CC CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.  
 CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE  
 CC FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AB010490; BAA31199.1; JOINED.  
 DR EMBL; AF190631; AAF01340.1; -.  
 DR EMBL; S72628; AAD14112.1; -.  
 DR PIR; S04459; OYHUR.  
 DR HSSP; 002846; 1AMU.  
 DR GeneW; HGNC:7943; NPRL.  
 DR MIM; 108960; -.  
 DR InterPro; IPR001828; ANP\_receptor.  
 DR InterPro; IPR000719; Euk\_kinase.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR InterPro; IPR001170; Ntpep\_receptor.  
 DR Pfam; PF00069; kinase; 1.  
 DR Pfam; PF00211; guanylate\_cyc; 1.  
 DR Pfam; PF01094; ANP\_receptor; 1.  
 DR PRINTS; PR00255; NATPEPTIDER.  
 DR ProDom; PD000001; Euk\_kinase; 1.  
 DR SMART; SM00044; CycC; 1.  
 DR PROSITE; PS00458; ANP\_RECEPTORS; 1.  
 DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE; PS00125; GUANYLATE\_CYCLASES\_2; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;  
 KW GMP synthesis; Signal.  
 FT SIGNAL 1..32  
 FT CHAIN 33..1061  
 FT DOMAIN 33..473  
 FT TRANSMEM 474..494  
 FT DOMAIN 495..1061  
 FT DOMAIN 528..805  
 FT DOMAIN 876..1006  
 FT DISULFD 92..118  
 FT DISULFD 196..245  
 FT DISULFD 455..455  
 FT DISULFD 464..464  
 FT CARBOHYD 34..34  
 FT CARBOHYD 45..45  
 FT CARBOHYD 212..212  
 FT CARBOHYD 338..338  
 FT CARBOHYD 379..379  
 FT CARBOHYD 386..386  
 FT CARBOHYD 427..427  
 SEQUENCE 1061 AA; 118918 MW; E6B5BD0FCA32F70D CRC64;  
 Query Match 14.98; Score 481.5; DB 1; Length 1061;  
 Best Local Similarity 27.78; Pred. No. 1.86-25;  
 Matches 197; Conservative 88; Mismatches 227; Indels 199; Gaps 31;  
 OY 33 EEGFIVRIITDDEKTDLVAAASKVN-----LNAGEI 66  
 DB 416 ENGAFRYVLTNTNGT-SQELVAVSGRKLWPLGTPPPDIPKCGFNEDPACNODHLSLEY 474  
 OY 67 LQMPGKM-----FFVF-----COESGYDILRVLSNVEFLQNLDAHDAHTYTP 113  
 DB 475 LALVGSLSLGLILVSPFYRKQMLEKELASLRVMEDEP-----SLERL----- 524  
 OY 114 GMRAPSFRCYDAENGK--GLILHYSEREGLDIVIGITKVAQOIHGTETDMKVIOQRN 171  
 DB 525 --RSAGSRLLTSGSGSNVSL-----TTEGGQVY--FAKAYYK--GNLVAVRVNRK 573  
 OY 172 EECHTQFLIEKSKEDFEEDLRFPEENGTOGSRISPYT-FCKAPPHIIFRDLYVT 230  
 DB 574 IEL--TRKVLFEKLMRDVONEHLTRVYAGCTDPNPICLITEYCPRGSLDILENESTTL 631  
 OY 231 -----OCGNAIYEVLPOLQPGN-CSLSVFSVLRPHIDISFGILSHINTV-----FVLR 279  
 DB 632 DMRFRYSLTNDIVYGMFLHNGAICS-----HNNKSSNVCVVDGRVYVK 675  
 OY 280 SKE-GLIDVEKLECEDELGTETISCLRLKGMIT--LPEADSLTFLCSPSV--MNTDDL 333  
 DB 676 ITDYGLESFRDLDE-----OGHTVYAKKIMTAPBELLRMAAPVRGSSQAGDV 722  
 OY 334 TRRGVLSDIPL-----HDATRDL-----VLLGRQ--FREYKTLQLEEL----- 372



Db 723 YSFGIILQEIARSGVFNHESGLSPKEIIEVTRNGEPPFRPSLALQSHLEELIMOR 782  
 Qy 373 -----TDRLQTLR-----ALEDER 387  
 Db 783 CWAEDPQERPPQQLRLFKFRNRENSNIIDNLRSMDQYANNLEELVEERQVLEEK 842  
 Qy 388 KTDPLLSVLPSPVANEIRKRPAPAKYDNTILFSGIVGFNACSHHAGEGMAKTY 447  
 Db 843 RAEALQLQILPHSAEQILKRGETQAEFDSVTTFPSIDVFTLALSAEST---PMQV 898  
 Qy 448 NLINDLYTRFDLTQSRKNPFYKVEYVDKMTWVSGLEDEPCIH---HARSICHTALM 503  
 Db 899 TLINDLYTCFDAVID---NEDYKVTETIGDAMVYSGLP---VRNRLHACEYARALAL 952  
 Qy 504 MEIAQOVQV---DGSVQITIGIHGEVVTGVIQGMPPYCLFGNTVNLSTRTTGEKG 560  
 Db 953 IDAVSFRIRHRPQDLRLRIGIHGVPACVAGYGLKMPRYCLFGDVTNVRMSNGEAL 1012  
 Qy 561 KINSEYTYRCIMSPENSDPQHLERHGVSKGKKEPQVWFLSKKNTGT 611  
 Db 1013 KIHLSSETKAVL---EEFGFELRLGDVEKMGKRVTYMLGERSST 1059

## RESULT 13

ANPA\_RAT STANDARD: PRT: 1057 AA.

AC P18910;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A) (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide A-type receptor).  
 GN NPRL  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA MEDLINE=89143770; PubMed=2563900;  
 RA Chinkers M., Garbers D.L., Chang M.S., Lowe D.G., Chin H., Goeddel D.V., Schulz S.;  
 RT "A membrane form of guanylate cyclase is an atrial natriuretic peptide receptor.";  
 RL Nature 338:78-83(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=91056089; PubMed=1978722;  
 RA Yamaguchi M., Rutledge L.J., Garbers D.L.;  
 RT "The primary structure of the rat guanylyl cyclase A/atrial natriuretic peptide receptor gene.";  
 RL J. Biol. Chem. 265:20414-20420(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=91352095; PubMed=1679239;  
 RA Duda T., Goraczniak R.M., Sharma R.K.;  
 RT "Site-directed mutational analysis of a membrane guanylate cyclase cDNA reveals the atrial natriuretic factor signaling site.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:7882-7886(1991).  
 CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE CYCLASE ACTIVITY ON BINDING OF ANP.  
 CC -1- CATALYTIC ACTIVITY: GTP = 3',5'-cyclic GMP + diphosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C) CIRCULATING WITHOUT A ROLE IN SIGNAL TRANSDUCTION.  
 CC -1- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC DOMAIN OF PROTEIN KINASES.  
 CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE FAMILY.

CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X14773; CAA32881.1; -  
 CC EMBL: J05677; AAA11200.1; -  
 CC EMBL: M74535; AAA11202.1; -  
 CC PIR: S03348; OYTR.  
 CC HSP: Q02846; IAWL.  
 DR InterPro: IPR001828; ANF\_receptor.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001054; G\_cyclase.  
 DR InterPro: IPR001170; Ntpep\_receptor.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF01094; guanylate\_cyc; 1.  
 DR Pfam: PF01094; guanylate\_cyc; 1.  
 DR PRINTS: PR00255; NTPPTIDER.  
 DR PRODOM: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00044; CYC; 1.  
 DR PROSITE: PS00458; ANF\_RECEPTORS; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KM Receptor; Transmembrane; Glycoprotein; phosphorylation; Lyase;  
 KM GMP synthetis; Signal;  
 FT SIGNAL 1 28  
 FT CHAIN 29 1057  
 FT DOMAIN 29 469  
 FT TRANSMEM 470 490  
 FT DOMAIN 491 1057  
 FT DOMAIN 524 801  
 FT DOMAIN 872 1002  
 FT DISULFID 88 114  
 FT DISULFID 192 241  
 FT DISULFID 451 451  
 FT DISULFID 460 460  
 FT CARBOHYD 41 41  
 FT CARBOHYD 208 208  
 FT CARBOHYD 334 334  
 FT CARBOHYD 375 375  
 FT CARBOHYD 382 382  
 FT CARBOHYD 423 423  
 FT CONFLICT 366 366  
 FT CONFLICT 392 392  
 SQ SEQUENCE 1057 AA; 118951 MW; 9EA9AE685AC05816 CRC64;  
 Query Match 14.5%; Score 468; DB 1; Length 1057;  
 Best Local Similarity 43.4%; Pred. No. 1.6e-24;  
 Matches 112; Conservative 38; Mismatches 82; Indels 26; Gaps 9;  
 Qy 361 EEXKTLQDELILRLQLRLALEDEKKTDLLSVLPSPVANEIRKRPAPAKRYDY 420  
 Db 817 EGYANMLE-ELVEERQAVL---EKKRAELLYQILPHSAEQILKRGETQAEFDSV 871  
 Qy 421 TLFSGIVGFNACSHHAGEGMAKTYNLINDLYTRFDLTQSRKNPFYKVEYVDKTY 480  
 Db 872 TLFSGIVGFNACSHHAGEGMAKTYNLINDLYTRFDLTQSRKNPFYKVEYVDKTY 924  
 Qy 481 TVSGLEPCPH---HARSICHTALMMEIAQOVQV---DGSVQITIGIHGEVVTGY 533  
 Db 925 VVSGLP---VRNQLHAREYARALALDLVRSFRIRHPQDLRLRIGIHGVPACVAGV 981  
 Qy 534 GQMPRYCLFGDVTNVRMSNGEALKIHLSSETKAVL---EEFGFELRLGDVEKMGKRV 593  
 Db 982 GLKMPRYCLFGDVTNVRMSNGEALKIHLSSETKAVL---EEFGFELRLGDVEKMGKRV 1037  
 Qy 594 GKKEPQVWFLSKKNTGT 611

Db 1038 GRGKVRFTWLLGRCST 1055

RESULT 14

ANPA\_MOUSE STANDARD; PRT: 1057 AA.

ID ANPA\_MOUSE

AC P18293;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)

DE (guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide A-type receptor).

GN NPRI OR NPRA.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J;

RX MEDLINE=90324219; PubMed=1973687;

RA Pandey K.N., Singh S.;

RT "Molecular cloning and expression of murine guanylate cyclase/atrial natriuretic factor receptor cDNA.";

RL J. Biol. Chem. 265:12342-12348(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=95139994; PubMed=7838126;

RA Schoenfeld J.R., Sehl P., Quan C., Burnier J.P., Lowe D.G.;

RT "Agonist selectivity for three species of natriuretic peptide receptor-A.";

RL Mol. Pharmacol. 47:172-180(1995).

CC -1 FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE CYCLASE ACTIVITY ON BINDING OF ANP.

CC -1 CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.

CC -1 SUBCELLULAR LOCATION: Type I membrane protein.

CC -1 MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C) WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.

CC -1 SIMILARITY: BELONGS TO ADENYLATE CYCLASE CLASS-4/GUANYLATE CYCLASE FAMILY

CC -1 SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: J05504; AAA37670.1; -

CC EMBL: L31937; AAA66945.1; -

CC PIR: A36568; OTMSAR.

CC PIR: A33088; A33088.

CC HSSP: Q02846; IAWL.

CC MGD: MGI:97371; NPRL.

CC InterPro: IPR001828; ANF\_receptor.

CC InterPro: IPR000719; Euk\_kinase.

CC InterPro: IPR001054; G\_cyclase.

CC InterPro: IPR001170; Ntpep\_receptorN.

CC Pfam: PF00069; pkinase.1.

CC Pfam: PF00211; guanylate\_cyc.1.

CC Pfam: PF01094; ANF\_receptor.1.

CC PRINTS: PR00255; NATPEPTIDER.

CC PRODOM: PD000001; Euk\_kinase.1.

CC SMART: SM00044; CYCC.1.

CC PROSITE: PS00458; ANF\_RECEPTORS.1.

CC PROSITE: PS00452; GUANYLATE\_CYCLASES.1; 1.

CC PROSITE: PS0125; GUANYLATE\_CYCLASES\_2; 1.

DR PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.

KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;

KW CGMP synthesis; Signal.

FT SIGNAL 1 28

FT CHAIN 29 1057

FT DOMAIN 29 469

FT TRANSMEM 470 490

FT DOMAIN 491 1057

FT DOMAIN 524 801

FT DOMAIN 872 1002

FT DISULFID 88 114

FT DISULFID 192 241

FT DISULFID 451 451

FT DISULFID 460 460

FT CARBOHYD 41 41

FT CARBOHYD 208 208

FT CARBOHYD 334 334

FT CARBOHYD 375 375

FT CARBOHYD 382 382

FT CARBOHYD 423 423

FT CONFLICT 3 3

FT CONFLICT 39 39

FT CONFLICT 122 122

FT CONFLICT 130 130

FT CONFLICT 285 285

FT CONFLICT 301 301

FT CONFLICT 404 405

FT CONFLICT 590 590

FT CONFLICT 652 652

FT CONFLICT 833 833

FT CONFLICT 958 958

FT CONFLICT 1044 1044

FT CONFLICT 1050 1050

FT CONFLICT 1055 1057

SO SEQUENCE 1057 AA; 119109 MW; 53A544FB2C8BE253 CMC64;

Query Match 14.48; Score 466; DB 1; Length 1057;

Best Local Similarity 28.48; Pred. No. 2; 1e-24;

Matches 190; Conservative 86; Mismatches 210; Indels 182; Gaps 32;

25 IRKEAOLDEE-GQFLVRIITDSDSKTYDL---VAASKYNL-----NAGEIILQMPKMPF 75

489 IYKMKLEKEVSELMKRVKREDQLQPSLEHNLASASRLTSLSGSNYSGLTTEQ FQ 547

76 VFQESGYD---TILRVIGSNV---REFLQNT---DALHDLATYQGRAPSPRCRD 124

548 VFAKTAIYVGNLVAVRKVRKRIELTRKVLFEKHKRDQVNEHL-TRFVG-----ACID 600

125 AERGRLLIHYEREGLDIVG-----IYTV-----AQQING--- 159

601 PP--NICILTEYCPKSLDILLENESITLDMWERYSLTNDIVGMLFLHNGAIGSHGNLK 658

160 -----TEIDMKVIOQRNECDHPTFLIEEKSEKREDYEDLRPEE---NG 202

659 SSNCVVDGKRVLKITTYGLESPDPPEEGHTLF-----AKLMTAPELIRASAPRAG 712

203 TQESRISPYTFCKAFPHIIFRDLYVTOG-----NATRYLPOLQPGNCSTL 250

713 SQAGDV-----YSGFIILQE--IALRSQVYVYEGLDLSPKEILERTVRGBQPP----- 758

251 LSVSLVRPHIDISFGIILSHINTVFVLSKBSGLVYEKLECDDELGTETISLRKQGM 310

759 -----FRPSMDIQ-----SHLEELQOL-----MORCWADDPERPFOQIRLAKRK 799

311 IYLPKADSLIFLCSPEVMNLDLITRGVYLSIDPLMDATRDVLLAGEQPREVKLQGE 370

800 EKNKNSNT-----LDNLISK-----MEQYANLLE-E 825

371 IITDLQLTLRALEDEKRTDPLLYSVLPSPVANELRHRPYPARKYDVTWTLIFSGIVG 430

826 IVEERTQAVL-----EKKRAEALVQILPHSVVAEQKRGETQVQAFAFDSVTIYFSDIVG 881

431 MAFCSHAGSGEMAKIVNLLNDLITRFDTLITDSKRNPFYKYETVGDKITMYSGLEPCTI 490

```

DB 882 TALSEST-----PMQVTTLLNDLYCFEDAVID---NFDYKETTIDDAIVVSGLP---V 931
OY 491 H-----HARSICHLADMMIEAGOVY---DGESVOITIGIHGEVNTVIGQMPRYCLE 543
DB 932 RRGQAHARVARMALALDAVSRFRHRHPRQBLRIRIGHPCVAGVGLKMPRYCLE 991
OY 544 GNTVNLTSRTETTGEGKGINVSEYTYRCIMSPENDPOFHLHRRGVPVSKGKKEPMQVNF 603
DB 992 GDTVNTASRMESENGEALRIHLSSETKAVL---EEDF-GFELRLGDEVEMKKGKVRVYWL 1047
OY 604 LSRKNTGT 611
DB 1048 LERGCST 1055

RESULT 15
ANPR_RAT STANDARD; PRT: 1047 AA.
ID ANPR_RAT STANDARD; PRT: 1047 AA.
AC P16067;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Atrial natriuretic peptide receptor B precursor (ANP-B) (ANPRB) (GC-B)
DE (Guanylate cyclase) (EC 4.6.1.2) (NPR-B) (Atrial natriuretic peptide
DE B-type receptor).
GN NPR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89376566; PubMed=2570641;
RA Schulz S., Singh S., Bellet R.A., Singh G., Tubb D.J., Chin H.,
RA Garbers D.L.;
RT "The primary structure of a plasma membrane guanylate cyclase
RT demonstrates diversity within this new receptor family."
RL Cell 58:1155-1162(1989).
CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC CYCLASE ACTIVITY ON BINDING OF ANP. SEEMS TO BE STIMULATED MORE
CC EFFECTIVELY BY BRAIN NATRIURETIC PEPTIDE (BNP) THAN BY ANP.
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC -1- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC DOMAIN OF PROTEIN KINASES.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M26896; AAA41205.1; -.
CC PIR: A33300; OYRDB.
CC HSSP: 002846; IAWL.
CC InterPro: IPR001828; ANP_receptor.
CC InterPro: IPR000719; Euk_kinase.
CC InterPro: IPR001054; G_cyclase.
CC InterPro: IPR001170; Ntpep_receptorN.
CC Pfam: PF00069; pkinaase; 1.
CC Pfam: PF00211; guanylate_cyc; 1.
CC Pfam: PF01094; ANP_receptor; 1.
CC PRINTS: PR00255; NATPEPTIDER.

```

```

DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS00458; ANP_RECEPTORS; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS00125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
KW GMP synthesis; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1047
FT DOMAIN 23 458
FT TRANSMEM 459 478
FT DOMAIN 479 1047
FT DOMAIN 513 786
FT DOMAIN 861 991
FT DISULFID 75 101
FT DISULFID 439 439
FT DISULFID 448 448
FT CARBOHYD 24 24
FT CARBOHYD 35 35
FT CARBOHYD 161 161
FT CARBOHYD 195 195
FT CARBOHYD 244 244
FT CARBOHYD 277 277
FT CARBOHYD 349 349
SQ SEQUENCE 1047 AA; 117126 MW; 5062C49228CC14A3 CRC64;

Query Match 14.1%; Score 454.5; DB 1; Length 1047;
Best Local Similarity 27.9%; Pred. No. 1.3e-23;
Matches 178; Conservative 79; Mismatches 238; Indels 143; Gaps 23;

OY 58 VLNLNGEILQMGKMEFYECQESGYDTILRVGSNVR-BELO--NLDALHDLATYPG 114
DB 458 IVALGTYGTFIRMGVSSFLFRKLMEKELASMLMRIMWELDPGNSDRH----- 508
OY 115 MRAPSRCTDAEGKG--LLHYYSEREGIADIYIGIKVAAQIIR--GTEIMKVIQOR 170
DB 509 -----KGASRLTSLRGSSYGSIMTAHGVYQJFANTGHEKGVVAIKHVNK 556
OY 171 NEBCDHTQFLIEEKSEKEEFYEDLDPEENGQESISPYTCAPPHIIFDRDLVYT 230
DB 557 RILTR-QVLEFKHNDVDF-NHLIRF-----IGACIDPPNIC-----IVTE 597
OY 231 QCGNAIYRVLPOLQPGNCILLSVLRPHID-----ISFHGISHTNV---EV 277
DB 598 YCRGSLQDLENDSTINDMFRYSILNDLVKGAFLHNSIISHGSKSNVCVSRFY 657
OY 278 LRSKE-GL-----LDVEKLECEDEL-----TGT-----ISCL 304
DB 658 LKTTDGLASFSTAEPPDSHALYAKKLTWAPELISGNPLPTTGMORADYSFAILLQEI 717
OY 305 RLKGOMIYLPEDASILFLCSPSYANLDDLFRRGVLSIDPLMDHATRLVILGE----- 357
DB 718 ALMSGPYLGLD-----LSPKETIYQVRNQRPRFPRSIDRTQINELVLMRCWAQDP 773
OY 358 -----QFRE-----YKLOEELIDRLQTLRALDEKRTDTL 393
DB 774 TERPDGQIGKFRFRKKEGSTILDLLRMBOYANNLEKIYEBRTQVLEKRRKRAEL 833
OY 394 LYVLPSPVANEIRKRPVAKRYDVTILFSIGVFNACSKHASSEGAKVNLNDL 433
DB 834 LYQILPISVAEQKRETVQAEAFDSVTIYFSIDVGTALSAEST--PMQVTTLLNDL 889
OY 454 YTRFDLTDSRKKNPFYKVTVDKWTYVSGLP-EPCIHARSICHLADMMIEAGOVY 512
DB 890 YTCFDALIID---NFDYKETTIDDAIVVSGLGRNGORAPETARALALDAVSSFR 946
OY 513 ---DGESVOITIGIHGEVNTVIGQMPRYCLEGNTVNLTSRTETTGEGKINSEYTY 569
DB 947 RHRPHQRLRIRIGHVHPVAGVGLKMPRYCLEGDTVNTASRMESENGALTKIHVSTTK 1006
OY 570 RCLMSPENDPOFHLHRRGVPVSKGKKEPMQVFLSK 607

```

Db 1007 DAL---DELGCFOLELRGQVEMKKGCMRTYWLIGER 1040

Search completed: June 27, 2003, 13:03:29  
Job time : 12.9305 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:00:58 ; Search time 31.2101 Seconds  
(without alignments)  
4086.601 Million cell updates/sec

Title: US-09-762-767a-4

Perfect score: 3231  
Sequence: 1 MYGFVNHALLELVIRNGPE.....QWFLSRKNTGTETKQDD 619

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: SP:archaea:\*  
2: SP:bacteria:\*  
3: SP:fungi:\*  
4: SP:human:\*  
5: SP:invertebrate:\*  
6: SP:mammal:\*  
7: SP:mnc:\*  
8: SP:organelle:\*  
9: SP:phage:\*  
10: SP:plant:\*  
11: SP:rodent:\*  
12: SP:virus:\*  
13: SP:vertebrate:\*  
14: SP:unclassified:\*  
15: SP:rvirus:\*  
16: SP:bacteriaph:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3198	99.0	620	11	O54865	O54865 mus musculus
2	2887	89.4	614	13	P79998	P79998 oryzias lat
3	2859	88.5	617	13	O90VY5	O90VY5 figu rubrip
4	1903.5	58.9	600	5	O77106	O77106 manduca sex
5	1883	58.3	758	5	O24086	O24086 drosophila
6	1883	57.9	787	5	O9VA09	O9VA09 drosophila
7	1870.5	57.9	604	5	O95NK5	O95NK5 hemocentrot
8	1761	54.5	649	5	O17010	O17010 anopheles g
9	1105	34.2	742	11	O91XJ7	O91XJ7 rattus norv
10	983	30.4	636	11	O92001	O92001 rattus norv
11	960.5	29.7	685	5	O9XTE0	O9XTE0 caenorhabd1
12	955	29.6	686	5	P92006	P92006 caenorhabd1
13	955	29.6	940	5	O76340	O76340 manduca sex
14	938	29.0	751	5	O02298	O02298 caenorhabd1
15	929.5	28.8	690	5	O9YFCS	O9YFCS drosophila
16	914	28.3	699	5	O17707	O17707 caenorhabd1

17	912.5	28.2	684	5	O9B180	O9B180 caenorhabd1
18	788	24.4	564	4	O9NNW8	O9NNW8 homo sapien
19	788	24.4	675	13	O90VW5	O90VW5 figu rubrip
20	786.5	24.3	690	4	O8TAH3	O8TAH3 homo sapien
21	783	24.2	678	13	O9PW12	O9PW12 oryzias lat
22	778	24.1	678	13	P79997	P79997 oryzias lat
23	778	24.1	691	11	O9DBQ3	O9DBQ3 mus musculu
24	778	24.1	691	11	O9ERL9	O9ERL9 mus musculu
25	757	23.4	699	5	O77105	O77105 manduca sex
26	729.5	22.6	752	5	O22301	O22301 caenorhabd1
27	726	22.5	676	5	O24085	O24085 drosophila
28	720	22.5	676	5	O95S04	O95S04 drosophila
29	704.5	21.8	669	5	O9VEU5	O9VEU5 drosophila
30	701.5	21.7	583	5	O44468	O44468 caenorhabd1
31	685	21.2	649	5	O9VED6	O9VED6 drosophila
32	646.5	20.0	947	5	P90895	P90895 caenorhabd1
33	565	17.5	108	6	O8SPV3	O8SPV3 ovls arles
34	493	15.3	220	11	O8R5L4	O8R5L4 rattus norv
35	490.5	15.2	1070	13	O93490	O93490 anguilla ja
36	478.5	14.8	1126	5	O9BPR0	O9BPR0 bombyx mori
37	472	14.6	1055	13	O98VU1	O98VU1 oryzias lat
38	471	14.6	1055	13	O9YGM3	O9YGM3 oryzias lat
39	470	14.5	1057	5	O18331	O18331 caenorhabd1
40	469.5	14.5	1052	13	O90YB7	O90YB7 rana catesb
41	467.5	14.5	1276	5	O878L6	O878L6 caenorhabd1
42	466	14.4	1057	11	O91X04	O91X04 mus musculu
43	464	14.4	1047	13	O9PMH0	O9PMH0 xenopus lae
44	458	14.2	1076	5	O9VFI9	O9VFI9 drosophila
45	456.5	14.1	632	5	O94429	O94429 ctoma intes

## ALIGNMENTS

RESULT 1	O54865	PRELIMINARY;	PRT;	620 AA.
ID	O54865			
AC	O54865;			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Soluble guanylate cyclase beta-1 subunit.			
GN	GUCY1B3 OR GC-S-BETA-1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LUNG;			
RX	STRAIN=C57BL/10;			
RA	PubMed=10984516;			
RA	Sharina I.G., Krumenacker J.S., Martin E., Murad F.;			
RT	"Genomic organization of alpha 1 and beta 1 subunits of the mammalian			
RT	soluble guanylyl cyclase genes";			
RT	Proc. Natl. Acad. Sci. U.S.A. 97:10878-10883(2000).			
DR	EMBL; AF020339; AAB94876.1; -;			
DR	EMBL; AF297083; AAG17447.1; -;			
DR	HSSP; P16068; IAWN.			
DR	MGI; MGI:1860604; Gucylb3.			
DR	InterPro: IPR001054; G_cyclase.			
DR	Pfam: PF00211; guanylate_cyc; 1.			
DR	SMART; SM00044; CYCC; 1.			
DR	PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.			
DR	PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.			
KW	LYASE.			
SO	SEQUENCE	620 AA;	70597 MW;	EED98A9F539F4EF9 CRC64;
Query Match		99.0%;	Score 3198;	DB 11; Length 620;



GN FRGS-BETA1.  
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Takifugu.  
 NCBI\_TaxID=31033;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Morinaga C., Yamamoto T., Moriya Y., Suzuki N.;  
 RT Identification of tandem organization of soluble guanylyl cyclase  
 RT alpha and beta subunit genes in the Japanese pufferfish (Fugu  
 rubripes) genome.  
 RL Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB062172; BAB60908.1; -  
 DR EMBL; AB062170; BAB60906.1; -  
 DR InterPro: IPR001054; G\_cyclase.  
 DR InterPro: IPR000834; zn\_carboxypept.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR PROSITE: PS00132; CARBOXYPEPT\_2N\_1; UNKNOWN\_1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; UNKNOWN\_1.  
 DR PROSITE: PS00125; GUANYLATE\_CYCLASES\_2; 1.  
 SQ SEQUENCE 617 AA; 70236 MW; 7B531B5896A06191 CRC64;

Query Match 88.5%; Score 2859; DB 13; Length 617;  
 Best Local Similarity 88.7%; Pred. No. 1.3e-207;  
 Matches 540; Conservative 41; Mismatches 28; Indels 0; Gaps 0;

1 MGFTVNALELVIRNGPEWEDIKKKAQDEGQFLVRIYDSKTYDLVAASKVYN 60  
 1 MGFTVNALELVIRNGPEWEDIKKKAQDEGQFLVRIYDSKTYDLVAASKVYN 60  
 1 MGFTVNALELVIRNGPEWEDIKKKAQDEGQFLVRIYDSKTYDLVAASKVYN 60  
 61 IDAGDILQFGKMFEEFCQESGYDTLIRVLSNVREFLQNLADHDLHGLITVPGMRAPSF 120  
 61 IDAGDILQFGKMFEEFCQESGYDTLIRVLSNVREFLQNLADHDLHGLITVPGMRAPSF 120  
 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAAQIHGTEIDMKYIQORNEBCHTOFL 180  
 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAAQIHGTEIDMKYIQORNEBCHTOFL 180  
 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAAQIHGTEIDMKYIQORNEBCHTOFL 180  
 181 IEKSKEDFEYDLDREFENGQESRISPYTFCKAFPHIITFDRLVLTQCGNAIYRYL 240  
 181 IEKSKEDFEYDLDREFENGQESRISPYTFCKAFPHIITFDRLVLTQCGNAIYRYL 240  
 181 IEKSKEDFEYDLDREFENGQESRISPYTFCKAFPHIITFDRLVLTQCGNAIYRYL 240  
 241 POLQPGNCSLVSFLVRPHIDISFGHLSHINTVFLRSKGLDVEKLECEDELGTGE 300  
 241 POLQPGNCSLVSFLVRPHIDISFGHLSHINTVFLRSKGLDVEKLECEDELGTGE 300  
 241 POLQPGNCSLVSFLVRPHIDISFGHLSHINTVFLRSKGLDVEKLECEDELGTGE 300  
 301 ISCLRKGOMIYLPKADSLIFLCSPSVMNLDLTRGLYSDIPLDATRDVLLGEORF 360  
 301 ISCLRKGOMIYLPKADSLIFLCSPSVMNLDLTRGLYSDIPLDATRDVLLGEORF 360  
 301 ISCLRKGOMIYLPKADSLIFLCSPSVMNLDLTRGLYSDIPLDATRDVLLGEORF 360  
 361 EEYKLTQELIETDRQLTLRALDEKKKTDTLLYSVLPPSVANELRHKRPVAKRYDNY 420  
 361 EEYKLTQELIETDRQLTLRALDEKKKTDTLLYSVLPPSVANELRHKRPVAKRYDNY 420  
 361 EEYKLTQELIETDRQLTLRALDEKKKTDTLLYSVLPPSVANELRHKRPVAKRYDNY 420  
 421 TILFSGIVGFNAFCSKHASEGAMKIVNLNDLYTRFDLTDSRKNPFYKYVTVGDKYM 480  
 421 TILFSGIVGFNAFCSKHASEGAMKIVNLNDLYTRFDLTDSRKNPFYKYVTVGDKYM 480  
 421 TILFSGIVGFNAFCSKHASEGAMKIVNLNDLYTRFDLTDSRKNPFYKYVTVGDKYM 480  
 481 TVSGILPEPCIHARSICHLADMMETAGOVQVDSGVQITIGITGEVYTVGVIGQMPRY 540  
 481 TVSGILPEPCIHARSICHLADMMETAGOVQVDSGVQITIGITGEVYTVGVIGQMPRY 540  
 481 TVSGILPEPCIHARSICHLADMMETAGOVQVDSGVQITIGITGEVYTVGVIGQMPRY 540  
 541 CLFGNTVNLSTRTETGEGKINVSYYTRCLMSPNSDPOFLERHGVSMKGKREPMO 600  
 541 CLFGNTVNLSTRTETGEGKINVSYYTRCLMSPNSDPOFLERHGVSMKGKREPMO 600  
 541 CLFGNTVNLSTRTETGEGKINVSYYTRCLMSPNSDPOFLERHGVSMKGKREPMO 600  
 601 VWFLSRKNP 609  
 601 VWFLSRKNP 609  
 601 VWFLSRKNP 609

RESULT 4  
 ID 077106 PRELIMINARY; PRT; 600 AA.  
 AC 077106;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Soluble guanylyl cyclase beta-1 subunit.  
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysta;  
 OC Spingidae; Spingidae; Spinginae; Manduca.  
 NCBI\_TaxID=7130;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-98409773; PubMed-9736646;  
 RA Nighorn A., Gibson N.J., Rivers D.M., Hildebrand J.G., Motton D.B.;  
 RT "The nitric oxide-cGMP pathway may mediate communication between  
 RT sensory afferents and projection neurons in the antennal lobe of  
 RT manduca sexta."  
 RL J. Neurosci. 18:7244-7255(1998).  
 DR EMBL; AF062751; AAC61264.1; -  
 DR HSSP; P16068; IAWN.  
 DR InterPro: IPR001054; G\_cyclase.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART; SM00044; CYCC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS00125; GUANYLATE\_CYCLASES\_2; 1.  
 DR LYASE.  
 SQ SEQUENCE 600 AA; 68099 MW; 7FD0136E6ADACB9 CRC64;

Query Match 58.9%; Score 1903.5; DB 5; Length 600;  
 Best Local Similarity 60.3%; Pred. No. 2.1e-135;  
 Matches 366; Conservative 94; Mismatches 138; Indels 9; Gaps 4;

1 MGFTVNALELVIRNGPEWEDIKKKAQDEGQFLVRIYDSKTYDLVAASKVYN 60  
 1 MGFTVNALELVIRNGPEWEDIKKKAQDEGQFLVRIYDSKTYDLVAASKVYN 60  
 1 MGFTVNALELVIRNGPEWEDIKKKAQDEGQFLVRIYDSKTYDLVAASKVYN 60  
 61 IDAGDILQFGKMFEEFCQESGYDTLIRVLSNVREFLQNLADHDLHGLITVPGMRAPSF 120  
 61 IDAGDILQFGKMFEEFCQESGYDTLIRVLSNVREFLQNLADHDLHGLITVPGMRAPSF 120  
 61 IDAGDILQFGKMFEEFCQESGYDTLIRVLSNVREFLQNLADHDLHGLITVPGMRAPSF 120  
 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAAQIHGTEIDMKYIQORNEBCHTOFL 180  
 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAAQIHGTEIDMKYIQORNEBCHTOFL 180  
 121 RCTDAEKKGGLILHYSEREGLODIYIGIKTVAAQIHGTEIDMKYIQORNEBCHTOFL 180  
 181 IEKSKEDFEYDLDREFENGQESRISPYTFCKAFPHIITFDRLVLTQCGNAIYRYL 240  
 181 IEKSKEDFEYDLDREFENGQESRISPYTFCKAFPHIITFDRLVLTQCGNAIYRYL 240  
 181 IEKSKEDFEYDLDREFENGQESRISPYTFCKAFPHIITFDRLVLTQCGNAIYRYL 240  
 241 POLQPGNCSLVSFLVRPHIDISFGHLSHINTVFLRSKGLDVEKLECEDELGTGE 300  
 241 POLQPGNCSLVSFLVRPHIDISFGHLSHINTVFLRSKGLDVEKLECEDELGTGE 300  
 241 POLQPGNCSLVSFLVRPHIDISFGHLSHINTVFLRSKGLDVEKLECEDELGTGE 300  
 301 ISCLRKGOMIYLPKADSLIFLCSPSVMNLDLTRGLYSDIPLDATRDVLLGEORF 360  
 301 ISCLRKGOMIYLPKADSLIFLCSPSVMNLDLTRGLYSDIPLDATRDVLLGEORF 360  
 301 ISCLRKGOMIYLPKADSLIFLCSPSVMNLDLTRGLYSDIPLDATRDVLLGEORF 360  
 361 EEYKLTQELIETDRQLTLRALDEKKKTDTLLYSVLPPSVANELRHKRPVAKRYDNY 420  
 361 EEYKLTQELIETDRQLTLRALDEKKKTDTLLYSVLPPSVANELRHKRPVAKRYDNY 420  
 361 EEYKLTQELIETDRQLTLRALDEKKKTDTLLYSVLPPSVANELRHKRPVAKRYDNY 420  
 421 TILFSGIVGFNAFCSKHASEGAMKIVNLNDLYTRFDLTDSRKNPFYKYVTVGDKYM 480  
 421 TILFSGIVGFNAFCSKHASEGAMKIVNLNDLYTRFDLTDSRKNPFYKYVTVGDKYM 480  
 421 TILFSGIVGFNAFCSKHASEGAMKIVNLNDLYTRFDLTDSRKNPFYKYVTVGDKYM 480  
 481 TVSGILPEPCIHARSICHLADMMETAGOVQVDSGVQITIGITGEVYTVGVIGQMPRY 540  
 481 TVSGILPEPCIHARSICHLADMMETAGOVQVDSGVQITIGITGEVYTVGVIGQMPRY 540  
 481 TVSGILPEPCIHARSICHLADMMETAGOVQVDSGVQITIGITGEVYTVGVIGQMPRY 540  
 541 CLFGNTVNLSTRTETGEGKINVSYYTRCLMSPNSDPOFLERHGVSMKGKREPMO 600  
 541 CLFGNTVNLSTRTETGEGKINVSYYTRCLMSPNSDPOFLERHGVSMKGKREPMO 600  
 541 CLFGNTVNLSTRTETGEGKINVSYYTRCLMSPNSDPOFLERHGVSMKGKREPMO 600  
 601 VWFLSRKNP 609  
 601 VWFLSRKNP 609  
 601 VWFLSRKNP 609

Db	532	CLEGNVNLISRCETIGVYGTINVSBDYINYLAMREDNDDEQELTYRCHYIMKKAEPQ	591
Qy	601	VWFLSRK	607
Db	592	TWFLTRK	598

## RESULT 5

AC	Q24086;	PRELIMINARY;	PRT;	758 AA.
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 20, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Soluble guanylyl cyclase beta subunit.			
GN	GYC-BETA-100B OR DGCB1 OR CG1470.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-OREGON-R.			
RX	MEDLINE-95318108; PubMed-7797526;			
RA	Shah S., Hyde D.R.;			
RT	"Two Drosophila genes that encode the alpha and beta subunits of the			
RT	brain soluble guanylyl cyclase."			
RL	J. Biol. Chem. 270:15368-15376(1995).			
DR	EMBL; U27123; AAA87941.1; -			
DR	HSSP; P16068; IAWN.			
DR	FlyBase: FBgn013973; Gyc-beta-100B.			
DR	InterPro: IPR001054; G_cyclase.			
DR	Pfam: PF00211; guanylate_cyc; 1.			
DR	SMART; SM00044; CYCC; 1.			
DR	PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.			
DR	PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.			
KM	Lysase.			
QO	SEQUENCE 758 AA; 84444 MW; FBIDC388142E9F09 CRC64;			

[illegible]

```

Db      : 407 --SRHQBRLRLKGOMATYIFETDRILFQOCYPSVMNLDLTFKKGLYISDPVPLHDAAROLYL 464
QY      : 355 LSGQFEBEKKTLOEELITDRQLRLRALDEBEKKTPTDLXSVLPSTVANTLRKKRPYA 414
        : 465 LSRKFAEAKYLRKNEMLTDKLDQDFRDLSEKOTDRLXSVLPKSVANMLRHQREVP 524
Db      : 415 KRDNATYILFSGVIGFNAFCSKHA5GEGAMKYVNLNLTLYTRFDLNDKRNPNVYVET 474
QY      : 525 KRDSYTLTMSGIVGQYCAANTPDGAMKYRKLNELTYVFDALDSKRNLVYVET 584
Db      : 475 VGDKYMTVSGLEPBCIHHARSICHLALDMELIAGOVYDGB5VOITIGIHTEGVTVIG 534
QY      : 585 VGDKYMAVSGLEPDCEDHAKCMARVALDDMDMAKIVKGN5NPVOTIGIH5GEVTVIG 644
Db      : 535 QMRPRCLCGNVNLT5RRTTGEKGKINVS5YTRCLMS5PNSDPQFHLHRBPV5MG 594
QY      : 645 NRPVRCLEGNVYNLT5RTEGTVGGRINVS5EYTRLLCMA1NOD5FHELYRGPVIMKG 704
Db      : 595 KRPMQVMEFLSRKNT---GTEET 614
QY      : 705 KRPMQVMEFLRATIS5ILGTSST 728
Db

```

## RESULT 6

ID	PRELIMINARY:	PRT:	787 AA.
GNVAA09			
AC	09VAA09		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAY-2002 (TREMBLrel. 20, Last annotation update)		
DE	GYCBETA100B protein.		
GN	GYC-BETA-100B OR GYCBETA100B OR CG1470.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	Nb taxid:7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BERKELEY;		
RC	MEDLINE-20196006; PubMed-10731132;		
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blaise J.R.G., Champe M., Pfeiffer B.D.,		
RA	Wan R.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Arril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
RA	Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck Y., Brockstein P., Brotler P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Deng A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Posler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merklow G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Pacלב J.M.,		
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Ribeiro K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Sidenklamas I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,		
RA	Styrkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		



RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AB03777; AAF57119.1; -.  
 DR HSP; P16068; IAN.  
 DR F1ybase; Fggn0013973; Gyc-beta-100B.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR Pfam; PF00211; guanylate\_cyc; 1.  
 DR SMART; SM00044; CYCG; 1.  
 DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 KW Lyase.  
 SQ SEQUENCE 787 AA; 86849 MW; ABD2F4E6D09C72B8 CRC64;  
  
 Query Match 58.3%; Score 1883; DB 5; Length 787;  
 Best Local Similarity 49.9%; Pred. No. 1,1e-133;  
 Matches 371; Conservative 102; Mismatches 125; Indels 146; Gaps 7;  
  
 QY 1 MYGFVNALELVIRNYGPEWEDIRKEAQLDEGQFLVRIYDSDKTYDVAASKVLN 60  
 DB 1 MYGFVNALELVIRNYGPEWEDIRKEAQLDEGQFLVRIYDSDKTYDVAASKVLN 60  
 QY 61 LNAGETLQMGKMFVFCQESGYDTLLRVLGSNVREFLQNLDAHDLATIPGMRAPE 120  
 DB 61 IPADLLELFQKTFEFQCGDSGYDKLYLGATPRFLQNLDAHDLATIPGMRAPE 120  
 QY 121 RCTDAEKGLILHYSEREGLODIYIGIKTVAQOIHGETIDMKVYOORNE----- 172  
 DB 121 RCT--EKDGLILHYSERGLEHIVIGVAKASLHGEVEIDIVKKEPEIDAEKE 178  
 QY 173 ----- 172  
 DB 179 RALAEHQQLLEDVAITTTGSAIVYVLAESSDAERNNNHNGSNNNGMANNNTVAVN 238  
 QY 173 -----BC-----DHQFLIE 182  
 DB 239 NNNDGQQLASETDPSIALSTCPILAQSPFCDDGDKQKCLRLKNKSDIERDHYQLIR 298  
 QY 183 E-----KESKEDFYDDLRFENGNGQESRISPTTCKAPPHIITRDRLVYQCGN 234  
 DB 299 EINVAAKSGVDAKKEVPDMEFL-----CEAPLISPATFCVPPFLMEDRQKIVQAGK 354  
 QY 235 AIVRYLPOLQPNCSLSVFSLYVRPHIDISFGIILSHINTVFLRSKGLDVEKLECEB 294  
 DB 355 AVSRVLPRAEENCSLIEVEAIRPLQNFENILSHINTIYVLOTROGAMS----- 406  
 QY 295 ELVTGEISCLRLKGOMIYLPBADSILFLCSPVNMIDLTGRGLYLSIDPLHDAITRDLY 354  
 DB 407 --SRHEQRLRLKGMXYIPEDIRLIFQCYPSVMNLDLTGKGLYSDVPLHDAARDLY 464  
 QY 355 LGEQREEEKLQOELEILDRLOTLRALADEKKTDITLLSYLPPSVANELRHKRPVA 414  
 DB 465 LSEKEAEKRLKRNEMLTDLQOQTFRDESEKQKDRLLYSVLPKSVANELRHQRPVP 524  
 QY 415 KRYDNTILFSGIVGFNFCSHAGSEGAMKIVNLLNDLYTRFDLTDSRKNPFYVYKE 474  
 DB 525 KRYDSTILFSGIVGQCYCAANTDPDGMKIVKMLNELTYFDALTDSCRNLNAYKVE 584  
 QY 475 VGDKTYVSGLDEPCTIHHARSICHLALDMMETIAGOVODGESVQITIGIHGEVYVIG 534  
 DB 585 VGDKTYVSGLDEPCHEDHAKMARVADMMAKKMGSPVQITIGIHSGEVYVIG 644  
 QY 535 QMNPRLCLEGNTVNLSTRTTGEKGINVSETTYTCLSPENSDFQFLHLEHGPVSMG 594  
 DB 645 NNVPRLCLEGNTVNLSTRTTGEVGRINVSSETTYTCLMAINODDSFLEHGPVIMG 704  
 QY 595 KKEPMQVFLSKRNT-----GTEET 614  
 DB 705 KPLPDMCWLTRATSSILGTSSIT 728

RESULT 7  
 Q95NK5 PRELIMINARY; PRT; 604 AA.  
 AC Q95NK5;  
 DT 01-DEC-2001 (TREMBLrel, 19, Created)  
 DT 01-DEC-2001 (TREMBLrel, 19, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel, 20, last annotation update)  
 DE Soluble guanylyl cyclase beta1.  
 GN HPGCS-BETA1.  
 OS Hemileictrous pulcherrimus (Sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidae; Echinodermata; Echinacea; Echinoida; Strongylocentrotidae;  
 OC Hemileictrous.  
 OX NCBI\_TaxId=7650;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tanabe Y., Suzuki N.;  
 RT "Genomic Structure and Expression of the Sea Urchin Soluble Guanylyl  
 RT Cyclase beta Subunit Genes."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB062387; BAB58877.1; -.  
 DR EMBL: AB062387; BAB58877.1; JOINED.  
 DR EMBL: AB062388; BAB58877.1; JOINED.  
 DR EMBL: AB062389; BAB58877.1; JOINED.  
 DR EMBL: AB062386; BAB56135.1; -.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR Pfam; PF00211; guanylate\_cyc; 1.  
 DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; UNKNOWN.1.  
 DR PROSITE; PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 SQ SEQUENCE 604 AA; 68198 MW; B1C79618705934D2 CRC64;  
  
 Query Match 57.9%; Score 1870.5; DB 5; Length 604;  
 Best Local Similarity 57.2%; Pred. No. 6.7e-133;  
 Matches 351; Conservative 112; Mismatches 126; Indels 25; Gaps 5;  
  
 QY 1 MYGFVNALELVIRNYGPEWEDIRKEAQLDEGQFLVRIYDSDKTYDVAASKVLN 60  
 DB 1 MYGFVNALELVIRNYGPEWEDIRKEAQLDEGQFLVRIYDSDKTYDVAASKVLN 60  
 QY 61 LNAGETLQMGKMFVFCQESGYDTLLRVLGSNVREFLQNLDAHDLATIPGMRAPE 120  
 DB 61 ISANDLLEAFGRMFVFCQESGYDTLLRVLGSNVREFLQNLDAHDLATIPGMRAPE 120  
 QY 121 RCTDAEKGLILHYSEREGLODIYIGIKTVAQOIHGETIDMKVYOORNE----- 180  
 DB 121 RCTRESDDALVILHYSERGLEHIVIGVRSYAKTLHGEVNEILKNGECDAVORA 180  
 QY 181 IEEKSEKEDFYDDLRFENGTO-----ESRISPTTCKAPPHIIFRDRLVYQCG 233  
 DB 181 IIEK-----VETAKIEKQARQNLALSKKPKISPTSLRILPHIMNANLVEQAG 232  
 QY 234 NATRYVLPOLQPNCSLSVFSLYVRPHIDISFGIILSHINTVAVLSKGLDVEKLEB 293  
 DB 233 NSIQRIIVPININNCMTDLFHTVRPHMETFFSILSHANTIVLTNGSVN-----P 286  
 QY 294 DELTGEISCLRLKGOMIYLPBADSILFLCSPVNMIDLTGRGLYLSIDPLHDAITRDLY 353  
 DB 287 NNPNRSIALAKKQGMILHVPESNVLLYLCSPHIVNDELROKELTSLDPLHDAITRDLY 346  
 QY 354 LGEQREEEKLQOELEILDRLOTLRALADEKKTDITLLSYLPPSVANELRHKRPVP 413  
 DB 347 LISERDEEEKLQOELEILDRLOTLRALADEKKTDITLLSYLPPSVANELRHRRVP 406  
 QY 414 AKRYDNTILFSGIVGFNFCSHAGSEGAMKIVNLLNDLYTRFDLTDSRKNPFYVYKE 473  
 DB 407 AKKECVTLTFSGISFGDFCRYS--HDAKIVSLNSVYTKFEDVAMN--NPDYVYKE 462  
 QY 474 TVGDKTYVSGLDEPCTIHHARSICHLALDMMETIAGOVODGESVQITIGIHGEVYVIG 533  
 DB 463 TVGDKTYVSGLDEPCHADHAKCIKAKMALEKELSDAVINIEGDIYITIGVSGEVYVIG 522

OY 534 GORPRCLGNYVNTLSRETTGKRVSEYTYRCLSPENSDFPHLEHNGVSMK 593  
 DB 523 GORPRCLGNYVNTLSRETTGKRVSEYTYRCLSPENSDFPHLEHNGVSMK 582  
 OY 594 GKRPQVWFLSRK 607  
 DB 583 GKRPQVWFLSRK 596

## RESULT 8

017010 PRELIMINARY: PRT: 649 AA.

AC 017010: 017007; 017008; 017009;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Soluble guanylyl cyclase beta subunit (Fragment).  
 GN GCSBFA.  
 OS Anopheles gambiae (African malaria mosquito).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
 OC Anopheles.  
 OX NCBI\_Taxid=7165;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-SUA;  
 RX MEDLINE=99124361; Pubmed=9927171;  
 RA Caccione A., Garcia B.A., Mathiopoulos K.D., Min G.S., Moriyama E.N.,  
 RA Powell J.R.;  
 RT "Characterization of the soluble guanylyl cyclase beta-subunit gene in  
 RT the mosquito Anopheles gambiae."  
 RL Insect Mol. Biol. 8:23-30(1999).  
 RN [2]  
 RP SEQUENCE OF 539-649 FROM N.A.  
 RC STRAIN-SUA, G3, GMMKG, AND MU;  
 RX MEDLINE=96400927; Pubmed=8807303;  
 RA Garcia B.A., Caccione A., Mathiopoulos K.D., Powell J.R.;  
 RT "Genetics of monophyly in African anopheline malaria vectors."  
 RL Genetics 143:1313-1320(1996).  
 DR EMBL; U42613; AAC47142.1; -  
 DR EMBL; AF017062; AAC47144.2; -  
 DR EMBL; U42612; AAC47141.1; -  
 DR EMBL; U42614; AAC47143.1; -  
 DR HSSP; P16068; 1AMN.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR Pfam; PF00211; guanylate\_cyc; 1.  
 DR SMART; SM00044; CYCC; 1.  
 DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE; PSS0125; GUANYLATE\_CYCLASES\_2; 2.  
 KM Lyase.  
 FT NON\_TER  
 SQ SEQUENCE 649 AA; 72514 MW; 376BAC0422059DB CRC64;

Query Match 54.58; Score 1761; DB 5; Length 649;  
 Best Local Similarity 52.08; Pred. No. 14e-124; Mismatches 116; Indels 108; Gaps 4;

OY 2 YGFVNHALLELVIRNYPPEWEDIRKEAOLDEGQFLVRIYDYSKYDVAASKVLNT 61  
 DB 1 YGFVNHALLELVIRNYPPEWEDIRKEAOLDEGQFLVRIYDYSKYDVAASKVLNT 60  
 OY 62 NAGEIILQMGKMFVFCQESGYDTILRVGNSVREFLQNDALHDLATYTPGRAPDSFR 121  
 DB 61 PADDIILEFQKTEFEFCQDSGYDKILQVLCATPDPFLQNDALHDLATYTPGRAPDSFR 120  
 OY 122 CTDAEKGKGLIHYYSRELDIVIGITVAQOIHGTIDMKVIOQR----- 171  
 DB 121 CT--ETNGQLVLYHSERPELHIVIGIVAVASKLGVVEIKIIRKGDVPEAPAKA 178  
 OY 172 ----- 171  
 DB 179 TAEPPVANTAAATPKARHSIPVYKSVPTISLDPANPELANGLCKRIILASTSSGCP 238

OY 172 -----EEDHVFLLIE---KSEEDFYEDLDREENGTOESR 207  
 DB 239 ARNGPEWQKPKPGPTAVERSBDHFOFLITEISGPTPRRSDENQDAVAVKAEPM 298  
 OY 208 ISPTPCKAPPEPHIIPDRDLVYQCGNAIRVLQPLQPGCSLISVSLVRPHIDISFHG 267  
 DB 299 ISPTPCKAPPEPHIIPDRDLVYQCGNAIRVLQPLQPGCSLISVSLVRPHIDISFHG 358  
 OY 268 ILSHINTEVYLRKNEGLIDVEKLECEDELTGTEISCLRLKGMIVLEADSIIFLCSPSV 327  
 DB 359 ILAHINTEVYLRKNEGLIDVEKLECEDELTGTEISCLRLKGMIVLEADSIIFLCSPSV 406  
 OY 328 MNIDDLTRKGLVSYDIPLDHATRDVLVLSQFPEEKVKGLEILDRQLTLRALEDEK 387  
 DB 407 MNIDDLTRKGLVSYDIPLDHATRDVLVLSQFPEEKVKGLEILDRQLTLRALEDEK 466  
 OY 388 KKTDTLLSVLPSPVANELHKKRPVPAKRRDNTIFSGIVGNACSKRASEGMAKIV 447  
 DB 467 OKTDRLLYSVLPKIVANELHKKRPVPAKRRDNTIFSGIVGNACSKRASEGMAKIV 526  
 OY 448 NLANDLTYRDTLDSRKPNFVYKVVETGDKYMTVSGLPKCIHARSICHLADMMEIA 507  
 DB 527 KMLNELTYRDTLDSRKPNFVYKVVETGDKYMTVSGLPKCIHARSICHLADMMEIA 586  
 OY 508 GQVQVDSGVQITIGITGCVTVGVIGQRPRLCGNYVNTLSRETTGKRVSEY 567  
 DB 587 KNYMGTEAMKRTIGISHGEVTVGVIGNRPRLCGNYVNTLSRETTGKRVSEY 646  
 OY 568 TYR 570  
 DB 647 TYR 649

## RESULT 9

091XJ7 PRELIMINARY: PRT: 742 AA.

AC 091XJ7: 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Soluble guanylyl cyclase beta 2 subunit.  
 GN GRCY1B2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-KIDNEY;  
 RX MEDLINE=21391941; Pubmed=11406623;  
 RA Koglin M., Vehse K., Budeus L., Scholz H., Behrends S.;  
 RT "Nitric Oxide Activates the beta 2 subunit of Soluble Guanylyl Cyclase  
 RT in the Absence of a Second Subunit."  
 RL J. Biol. Chem. 276:30737-30743(2001).  
 DR EMBL; AF004153; AAF86581.1; -  
 DR InterPro; IPR001054; G\_cyclase.  
 DR Pfam; PF00211; guanylate\_cyc; 1.  
 DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; UNKNOWN\_1.  
 DR PROSITE; PSS0125; GUANYLATE\_CYCLASES\_2; 1.  
 SQ SEQUENCE 742 AA; 83050 MW; 41CE148B15E244E4 CRC64;

Query Match 34.28; Score 1105; DB 11; Length 742;  
 Best Local Similarity 36.98; Pred. No. 6.6e-75; Mismatches 221; Indels 88; Gaps 13;

OY 1 MYGFVNHALLELVIRNYPPEWEDIRKEAOLDEGQFLVRIYDYSKYDVAASKVLNT 60  
 DB 1 MYGFVNHALLELVIRNYPPEWEDIRKEAOLDEGQFLVRIYDYSKYDVAASKVLNT 58  
 OY 61 LNAEIIQMGKMFVFCQESGYDTILRVGNSVREFLQNDALHDLATYTPGRAPDSFR 120  
 DB 61 LNAEIIQMGKMFVFCQESGYDTILRVGNSVREFLQNDALHDLATYTPGRAPDSFR 120

```

Db 59 VSMELILKFGYEFKFCMSGYDRMLRPLGNGLEFIINLDALHSYALSTOEMNAPSF 118
Qy 121 RCTDAEKGLILHYYSEREGLODVIIGITVAQIHSTEDIMKVIQOORNEC----- 174
Db 119 RVEEGADG-AMLLHYYSRDRHGLCHIVPGIIEAVAKFFPTDVAMSLDM-NEEVERTGK 176
Qy 175 DHTOFLIEKESKE-----EDFYED----- 194
Db 177 EHAVFLVVOAKHROIRGAKASRPOGSEDSQADQALQTLIRKERYLINIPVCPGKSHS 236
Qy 195 -----LDREENGTOESRISPYTFCKAPFHIIIFRDLYVTQCGNATYRL 240
Db 237 TAVRASVLFKGPPLRDFQPYPERLWEEVFCDAFPFHIVFDALRVKQAGVNIQKY 296
Qy 241 POLQGNCSLSVSLVPRPHIDISFHLISHTVFLRSKGLDVEKLECEDELGTGE 300
Db 297 PCLIKQFALDEYFSIIHPQVTENISSICKFINSQFVLKTRKEMPKAR-----KS 347
Qy 301 ISCLALKGMITLPRADSLIFLCSPSVMDLITRGLYLSDIPLDATRDVLLGEOR 360
Db 348 QPMLLRGMIMESLRCMIFMCSFVNSISQLEESKMLSDIAPHDTRDILILNQOYL 407
Qy 361 EEKTLQOELEILDRLOTLRALDEKKTDTLLYSVLPVSANELRHRKRPVAKRYDNY 420
Db 408 AMELISCOLLEKKEKELRVLSNHLAIKKKTETLLYAMLEPHVANOLKGRKVAAGEFTJC 467
Qy 421 TILFSGIVGFNAFCSKSHASGEGAMKIVNLDLYTRFDLTDSRKNPFYKVTVDKYM 480
Db 468 TILFSDVVTFTNIC--AAFC-PIQIVNMLSMYSKFDRLTSVHD--YKVTETIDAM 520
Qy 481 TVSGLEPCHIHARSICHLADMMELAGOV--VDGESVQITIGITGEVTVGIGORNP 538
Db 521 YVGAVPVVESHQAORVANFALGMRSKAKVNPVGEPIQIRVGITGVLAGVVDKRP 580
Qy 539 RYCLFGNVTNLSRTETGKGINVSEYTRCLMSPENDPOFHLHGRPVSMKCKPE 598
Db 581 RYCLGDTVNTASRMSHGLPSKVLHSPAHRAL-----KNGKEIYVRGEIEVKGK- 634
Qy 599 MOVFLSRKNTGTEE 613
Db 635 MTFYFLIONLNTATED 649

```

RESULT 10

Q92001

ID Q92001 PRELIMINARY; PRT; 636 AA.

```

AC Q92001;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Soluble guanylate cyclase beta 2b.
OS Rattus norvegicus (Rat).
OC Mammalia; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RA Okamoto H., Asakawa T.;
RT *cDNA cloning and expression analysis of novel variant of rat sgc
RT beta2 subunit.*;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB058888; BAB68564.1; -
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PFO0211; guanylate_cyc; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASRS_1; UNKNOWN_1.
DR PROSITE; PS50125; GUANYLATE_CYCLASRS_2; 1.
SQ SEQUENCE 636 AA; 71642 MW; 931DBD8970E0890 CRC64;

```

Query Match 30.4%; Score 983; DB 11; Length 636;  
 Best Local Similarity 36.7%; Pred. No. 8.9e-66;  
 Matches 224; Conservative 104; Mismatches 196; Indels 86; Gaps 12;

```

Qy 66 ILQMGKMEFVFCQESGDTITLVIGSNVREFLONLDALHDLATIPQMRAPSFRCCTDA 125
Db 4 ILKLFGEYEFKFCMSGYDRMLRPLGNGLEFIINLDALHSYALSTOEMNAPSFVEEG 63
Qy 126 EKKGILHYYSEREGLODVIIGITVAQIHSTEDIMKVIQOORNEC-----DHTOP 179
Db 64 ADG-AMLLHYYSRDRHGLCHIVPGIIEAVAKDFPTDVAMSLDM-NEEVERTGKHEVVF 121
Qy 180 LIEKESKE-----EDFYED----- 194
Db 122 LVYQKHRDROIRGAKASRPOGSEDSQADQALQTLIRKERYLINIPVCPGKSHS 181
Qy 195 -----LDREENGTOESRISPYTFCKAPFHIIIFRDLYVTQCGNATYRLPOLP 245
Db 182 SVLFKGPPLRDFQPYPERLWEEVFCDAFPFHIVFDALRVKQAGVNIQKYVGILT 241
Qy 246 GNCISLSVSLVPRPHIDISFHLISHTVFLRSKGLDVEKLECEDELGTETISCLR 305
Db 242 QKRALDEYFSIIHPQVTENISSICKFINSQFVLKTRKEMPKAR-----KSOPMLK 292
Qy 306 LKGMITLPRADSLIFLCSPSVMDLITRGLYLSDIPLDATRDVLLGEORREYKL 365
Db 293 LRQGMIMESLRCMIFMCSFVNSISQLEESKMLSDIAPHDTRDILILNQOYLAEMEL 352
Qy 366 TQLELEILDRLOTLRALDEKKTDTLLYSVLPVSANELRHRKRPVAKRYDNTILFS 425
Db 353 SCQLEKKEKELRVLSNHLAIKKKTETLLYAMLEPHVANOLKGRKVAAGEFTCTILRS 412
Qy 426 GIYGFNAFCSKSHASGEGAMKIVNLDLYTRFDLTDSRKNPFYKVTVDKYMVSGI 485
Db 413 DVVTFFTNIC--AAFC-PIQIVNMLSMYSKFDRLTSVHD--YKVTETIDAMVVGAV 465
Qy 486 PERCHIHARSICHLADMMELAGOV--VDGESVQITIGITGEVTVGIGORNPXYCL 543
Db 466 PUVVESHQAORVANFALGMRSKAKVNPVGEPIQIRVGITGVLAGVVDKRPYCLF 525
Qy 544 GNTVNTLSRTETGKGINVSEYTRCLMSPENDPOFHLHGRPVSMKCKPEMQVNF 603
Db 526 GDTVNTASRMSHGLPSKVLHSPAHRAL-----KNGKEIYVRGEIEVKGK-MTFYF 579
Qy 604 LSRKNTGTEE 613
Db 580 LIONLNTATED 589

```

RESULT 11

Q9XTE0

ID Q9XTE0 PRELIMINARY; PRT; 685 AA.

```

AC Q9XTE0;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE C46E1.2 protein.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodieridae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Boufford J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Dublin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kerhaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Ritken L., Roopra A., Saunders D., Shownkeen R.,
RA Slaton N., Smith A., Sonhammer E., Staden R., Sulston J.,

```

RA Thierri-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sprat J., Wohlman P.,  
 RT 2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.  
 RL Nature 368:32-38(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RA Leonard N.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z99942; CAB17073.1; -  
 DR EMBL: AL008867; CAB17073.1; JOINED.  
 DR EMBL: AL008867; CAB15512.1; -  
 DR EMBL: Z99942; CAB15512.1; JOINED.  
 DR HSSP: Q02846; 1MW.  
 DR InterPro: IPR001054; G\_cyclase.  
 DR InterPro: IPR001230; Prenyl\_site.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART: SM00044; CYC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS00125; GUANYLATE\_CYCLASES\_2; 1.  
 DR PROSITE: PS00294; PRENYLATION; UNKNOWN\_1.  
 KW Lyase.  
 KW SEQUENCE 685 AA; 77871 MW; 68064419692A7BBE CRC64;

Query Match 29.7%; Score 960.5; DB 5; Length 685;  
 Best Local Similarity 34.2%; Pred. No. 4.9e-64;  
 Matches 220; Conservative 131; Mismatches 234; Indels 59; Gaps 17;

QY 2 YGFVNAHLELVIRNNGPEWEDIKKEAQDDEGQFLVRIYDDSKTYDLVAASKVLN 61  
 DB 12 FGFHESIRQLMRTVEGEAFWMSKVLERAGFEAKENINNYSDADPSLDVANSVLKV 71  
 QY 62 NAGEIILOMGKMFVFCQESGYDITLAVLSNREPLQNDALH---DHATY-PMRA 117  
 DB 72 TRCQVWMEYCCFLQYMETGMDLIRMSPNLKGFLNDLSLHYFDHV--VYKANLRG 129  
 QY 118 PSFRCTDAEKGKGLIHVYSEREGIDIVGIITVAQOIHGEIDMKV--OQRNEC- 174  
 DB 130 PSRCEEDNPDGT-TLHLYTGRGLYPIVAGVLEAKRKVFKLDVSTITGRQSVOMA 188  
 QY 175 -----DHQFLIE--EKSEKEDFEYEDLREENGTOESRISPYFCAFPFIIFDR 225  
 DB 189 TGERIEHVLFLVKTLMQDSNEALGTAVVQHSNN--YKIRLHMDFISFPFHVVQ 246  
 QY 226 DLYVTCGNALYKVLPO-LQPGNCSLSVSLVRPHIDISFGLSHINTVFLRSKGL 284  
 DB 247 DCRIVQGRRLYHNPRLDLSVGTPLMRIFEVTRPQIPDLFDSICNFIVNVLQVKTTP 306  
 QY 285 LDVEK-----LECEDE-----LGTETISCLRLKGOMTYLPEADSLFLGSPV 327  
 DB 307 NEQVNRANKRAAQAIESEMLYEDNKGALSLSQSHKLKGOMMLMSGGHIMILCSPY 366  
 QY 328 MNIDLTRGLYSLDPLDADTDVLLGEOFREYKLTQELITLRLQTLTALDEK 387  
 DB 367 TSIPLOYGLRLTAMPIHDPTRDLILNQRLSDVEMNLQLEANNQLEMAADLVEK 426  
 QY 388 KKTDTLLSYLSPSVANLHKKRPYAKRYDNTILFSGIVGF--AFCSKHSAGSGAM 444  
 DB 427 GKTDALREMLPPSVAQQLQGLSVEARETEAVMTDVPYTFQIYPLCT-----PK 479  
 QY 445 KIYNLNDLYTRDPLDLSRKNPFVYKVEYVGDKYMTVSGLPERCIIHARSICHALDM 504  
 DB 480 DIYHLNELFTKFDRLIGLOK---AKVETVGSYSKVGSIPLVDYDHCEYICHALGMV 536  
 QY 505 ELAGQV--QVDSGVOITIGIHGEVTVGVIGOMPRYCLFGMTVNLTSRTETGEGKI 562  
 DB 537 MEARTVCDPIITNPLIRAGIHSGPVAGVYAKMPRYCLFGDVTNNSRSHESPIGRI 596  
 QY 563 NSEYTYRCLMSPENSDPOHLEHGRVYSKKGKKEPQVFLSR 606  
 DB 597 HCSENAKCC---AESTGREFEPRGRVQIKGGE-MNTYFLLR 635

RESULT 12

ID P92006 PRELIMINARY; PRT; 686 AA.  
 AC P92006;  
 DT 01-MAY-1997 (Tremblrel. 03, Created)  
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE M04G12.3 protein.  
 GN M04G12.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI-TaxID=6237;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sims M.A.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C. elegans: A platform for  
 RT investigating biology";  
 RL Science 282:2012-2018(1998).  
 DR EMBL: Z81103; CAB03210.1; -  
 DR HSSP: P16068; 1MW.  
 DR InterPro: IPR001054; G\_cyclase.  
 DR InterPro: IPR001230; Prenyl\_site.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR SMART: SM00044; CYC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS00294; PRENYLATION; UNKNOWN\_1.  
 DR SEQUENCE 686 AA; 78588 MW; 65AE352DB53CBA57 CRC64;

Query Match 29.6%; Score 955; DB 5; Length 686;  
 Best Local Similarity 34.2%; Pred. No. 1.3e-63;  
 Matches 230; Conservative 126; Mismatches 235; Indels 82; Gaps 20;

QY 1 MYGFVNAHLELVIRNNGPEWEDIKKEAQDDEGQFLVRIYDDSKTYDLVAASKVLN 60  
 DB 1 MGFHESIRQLMRTVEGEAFWMSKVLERAGFEAKENINNYSDADPSLDVANSVLKV 60  
 QY 61 LNAGEIILOMGKMFVFCQESGYDITLAVLSNREPLQNDALH---DHATY-PMRA 116  
 DB 61 VTKQIWMYMGGLITYSMEIGWELVRMSPNLKGFLNDLSLHYFDHV--VYKANLR 118  
 QY 117 APSFRCTDAEKGKGLIHVYSEREGIDIVGIITVAQOIHGEIDMKVIOQRNEC-- 174  
 DB 119 GPSRCEENPDGT-TLHLYTGRGLYPIVAGVLEAKRKVFKLDVSTITGRQSVH 176  
 QY 175 -----DHQFLIEK-----SKEDFEYEDLREENGTOESRISPY 211  
 DB 177 MNNGEREVEHVLIFLKVEEPRRSDSTTSALTYSVPDGEIYD-----DNLKVSQ 229  
 QY 212 TFCAPFPHIIFBDLVVTCGNALYKVLPO--LQPGNCSLSVSLVRPHIDISFGL 269  
 DB 230 DFSRALPYHVLDESCLVCGDELYHINRELQPET-PILKIFENRQIPLDENIC 288  
 QY 270 SHINTVFLRSKGL-----DYKLECEDELGTETISCLRLKGOMTYL 313  
 DB 289 NFYNAYVVLQVKTSPFLAKKHMNMTKERQGEVAMEEVEESNLTQCHKLKGOMML 348  
 QY 314 PEADSLFLCSPSVANIDLTRGLYSLDPLDADTDVLLGEOFREYKLTQELITL 373  
 DB 349 STKHHIYLSYVTSINELMOGRMLTAMPLDADTDVLLNQRLSDVEMNLQLEANN 408  
 QY 374 DRLQTLRALDEKKKKTDTLLSYLSPSVANLHKKRPYAKRYDNTILFSGIVGNF 433  
 DB 409 EOLEYTMHLEVEERQKDTSLKQMLPKRIAKQLISGHEHPECE-ATVFCOLPAFOI 467  
 QY 434 ---CSKHSAGSAMKIYNLNDLYTRDPLDLSRKNPFVYKVEYVGDKYMTVSGLPERCI 490  
 DB 468 IYVQ-----PKNTYKILNEVFFKIDRLIVLRG---YKAVETVSDSTWTSIGIDYTS 517

```

OY 491 HHARSICHLALDMMELIAGOVQ--VDGESVQITIGIHTEGVTVIGIGORMPRYCLFGNTVN 548
DB 518 EHAEMCHVALGMMEARSVMDPVNKTPLRLIGHSGITLIGVYGTMKPRCLGEGTAT 577
OY 549 LTRSETTGEGKINVSSETTYRCLMSPENDPOFHLHNGPVSMMGKKEPMQWFLSR-- 606
DB 578 LASQESLGVAGKICQSSWTY---SKAMETGFEPSPRGIRNVRGRD-VETFTLMRSL 632
OY 607 KNTGTEETKODDD 619
DB 633 KKSITWEITDHERD 645

RESULT 13
ID 076340 PRELIMINARY; PRT: 940 AA.
AC 076340:
DB 01-NOV-1998 (TREMBLrel. 08, Created)
DB 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DB 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Soluble guanylyl cyclase beta-3.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditytola;
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
OX NCBI_Taxid=7130;
RN [1]
RP MEDLINE=99107914; PubMed=9891024;
RA Nigam A., Byrnes K.A., Morton D.B.;
RT Identification and characterization of a novel beta subunit of
RT soluble guanylyl cyclase that is active in the absence of a second
RT subunit and is relatively insensitive to nitric oxide."
RL J. Biol. Chem. 274:2525-2531(1999).
DR EMBL: AF064514; AAD09836.1;
DR HSSP: Q02846; IAWT.
DR InterPro: IPR001525; C5_DNA_meth.
DR InterPro: IPR001054; G_cyclase.
DR InterPro: IPR001230; Prenyl_site.
DR Pfam: PF00211; guanylate_cyc; 1.
DR SMART: SM00044; CYCC; 1.
DR PROSITE: PS00095; C5_MTASE_2; UNKNOWN_1.
DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.
SQ SEQUENCE 940 AA; 106393 MW; 0895688667BD0F6 CRC64;

Query Match 29.6%; Score 955; DB 5; Length 940;
Best Local Similarity 35.1%; Pred. No. 2e-63;
Matches 221; Conservative 123; Mismatches 234; Indels 52; Gaps 13;

OY 1 MTGFVNAHLELLIYRNGPEVWEDIKKEAQLDEGQFLVRIITDYSKTYDLVAAASKVYN 60
DB 1 MTGLLENNAEYIRTYGERBEDIRKQGV-EQPSFHYOYPPENLIRLAKKQAEVLG 59
OY 61 LNAGETLQMGKMFVFCQESGYDTLLRVLSNVREFLQNLDAHDHLLATTPGMRAPSF 120
DB 60 ITEREFQMGVYFVGFVSQYGYDRLSVLGRHMRDLNGLNHLREYLFKSPRMAPSF 119
OY 121 RCTDEKKGGLLHYHSEREGQDIYIGIKTYAQQIHTEIDMKYIQO----- 170
DB 120 IC-EMETROGLTLHRSKRGRGVYAMGQIRVAFRHYKEMKIELRELLFDVYVTF 178
OY 171 NEECHOTFLIEKSKSEDEYEDLRFENGTOESRIPTYECKAFPHIIFDRDLVYT 230
DB 179 QLTFFNRATFLASLMTREE-----KHLPIASVLELFPPCIYFGSMVVR 225
OY 231 QCGNATYRVLPOLQPEKNSLSVSLVRPHIDISFGHLSHNTVFLRSKSGLDV--- 287
DB 226 SIGNSLMTVLPDLV-GKKITWFDLVRPLIAFKFQITLNRNNIFELVYVEAVMEKAP 283
OY 288 -----ELTCEDELTGTEISCLRLKGMITLPEADSLTFLCSPSVNMLDLTLRGLYLSOI 343

```

```

DB 284 DKRNELLRLSDFTDCTTEKNLRLKGMITYMNMRRMNTLGTPEMDLSALVSTGYINDL 343
OY 344 PLHDAVRDLVLGEOFREERYKLTQELE-ILTRLOLTALBDEKKTPDILLYSLPSV 402
DB 344 SMHDSRLMLAGTQOSVBLKIALDOEOKSKLSEMRKLDENKRRIDELLYOMIPKOV 403
OY 403 ANELRH-KRPV-PARRVDNTILFSGIVGFNAFCSHASGEGAMKIVMLNDLYRFPDL 460
DB 404 ADLRNGENPIDTCEMFDSVSLFSDVYTFELISRT---PMEVSLNMYSLFPTL 459
OY 461 TDSKKNPFYKVEYVDKTYVSGUPERCIIHARSICHLALDMMELIAGOVQ--VDGESVQ 518
DB 460 TERNR---YKVEITGDATMYVSGAPEKEDNAEYCDALMDVATIDKDPSTGSHLS 516
OY 519 ITIGIHTEGVTVIGIGORMPRYCLFGNTVNLTSTETGEGKINVSSETTYRCLMSPENS 578
DB 517 IRGVHSGAVVAGIYGLKMPKRCLEGDSVNTASRHESTSEAKRIHISQTTQELL----- 570
OY 579 DPQFHLHNGPVSMMGKKEPMQWFLSRKN 608
DB 571 SPSTWYTERGEIQVGKGMKTYWLEGRES 600

RESULT 14
ID 002298 PRELIMINARY; PRT: 751 AA.
AC 002298:
DB 01-JUL-1997 (TREMBLrel. 04, Created)
DB 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DB 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE T04D3.4 protein.
GN T04D3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Kershaw J.K.;
RT Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RL none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: Z81114; CAB03288.1; -.
DR HSSP: P19687; IAWN.
DR InterPro: IPR001054; G_cyclase.
DR InterPro: IPR001230; Prenyl_site.
DR Pfam: PF00211; guanylate_cyc; 1.
DR SMART: SM00044; CYCC; 1.
DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.
SQ SEQUENCE 751 AA; 85380 MW; BC864A01A588CE9 CRC64;

Query Match 29.0%; Score 936; DB 5; Length 751;
Best Local Similarity 35.0%; Pred. No. 2.8e-62;
Matches 220; Conservative 125; Mismatches 236; Indels 48; Gaps 15;

OY 1 MTGFVNAHLELLIYRNGPEVWEDIKKEAQLDEGQFLVRIITDYSKTYDLVAAASKVYN 60
DB 1 MTGWHIESRQVLTAKRYGDIWEKIVHMSKFEIGTSETLAHYINDEETLRVNSMANYVG 60
OY 61 LNAGETLQMGKMFVFCQESGYDTLLRVLSNVREFLQNLDAH--DHLATY-PGMR 116
DB 61 IPLEIWEAYGFLIDFTMETGMDLRLAMPADLBEGFLDSLSLHYFIDHV--VYTKLR 118
OY 117 ABSFRTDMEKKGGLLHYHSEREGQDIYIGIKTYAQQIHTEIDMKYIQO----- 175

```

Db	119	GPSPRC-DVQADGCTLLHLYSKRSGLYPIVKGVREVARRIYDETVAKVQERQEHDA	177
Qy	176	-----HQFLIEE-----SKEDFYDDDRFBNGTOESRISPYTFCAPPHI	221
Db	178	FTEHVFVFITQIEMANSSTOPKSISSKASQIDL-----STGIYE--ISSDSFLAPYHI	231
Qy	222	IFPRDLVVTQCGNAYIRVLPQIQPGNCSLSYFSLYVRHIDISFGILSHINTVFVRSK	281
Db	232	CFPBDLFEHFGNFIKTFPPNARQETRTDLELVHEVPEPSESIKYVNSLFEVRK	291
Qy	282	EGLIDVEKLECEDELCTGCTISCLRKGQIVYLPREADSILFLCSPVNMJDLPTRGYLS	341
Db	292	-GLGDIY-----HNANDEKTYLLKGSVYFIDEBKTYLLYMSVATYVRELIRNLHS	344
Qy	342	DIPLDATRDVLVLLGQFREFEKKIQDEILTDRLQTLRALEDDEKKTDTLLYSVLPS	401
Db	345	DMORHDTBIVIMANOSRMSOVELNRTTEETTKTKKMAOELEIEKQTDDELCELPAS	404
Qy	402	VANELRHRKRVPAKRYDNTYILFSGYGVNARCASHASEGAMKIVNLTNLYRPFIT	461
Db	405	VADLSISGRKAMDKKEPADCTLFPTDLYVTNINCACTPYD---VTLTNDLYRFDLV	466
Qy	462	DSKRNPVYKVEYTGDKRYMTVSGLEPFCITHNARSICHLALDMELTAGV--QVDSGYOI	519
Db	461	GLMD--AYKEVETIGDAYIMYGVGPERCENAHAEVRLNISIGLMESKLYLPIIHKPIKI	511
Qy	520	TIGIHGEVYTYIGQMRPRCLFGYTNVLISRTETTGKKGKINSEYTYCLMSPEND	579
Db	518	RLGVHCGPVVAGVYGKIMPRYICFGDTYVANKMESNGIOCKIHSE--TGKILGLKAN	574
Qy	580	POFLEHRGVPVSMKGKKEPMQWFLSRKN	608
Db	575	PSYVFDRCNGTEIRG-KGMATYTFELERND	602
RESULT 15			
Q9VECS			
AC	Q9VECS	PRELIMINARY:	PRT: 690 AA.
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	CG4154 protein.		
GN	CG4154.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxId=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blaisey R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abriil J.F., Aghayani A., An H.-T., Andrews-Pfannkoch C., Baldwin D.,		
RA	Ballew F.M., Basu A., Bendale J., Bayraktiroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadelstein P., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasner K.,		
RA	Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris M.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Mei M.-H., Ibeyan C.,		
RA	Jatani M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		

RA Kimmel B.E., Kodliha C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzo L., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Relnett R., Remington K., Saunders R.D., Schaefer F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003707; AAF55135.1; -.  
 DR HSSP; Q02846; IAWL.  
 DR FlyBase; FBgn0038295; CG4154.  
 DR InterPro; IPR001525; G5\_DNA\_meth.  
 DR InterPro; IPR001054; G\_cyclase.  
 DR Pfam; PF00211; guanylate\_cyc; 1.  
 DR SMART; SM00044; CYCC; 1.  
 DR PROSITE; PS00095; G5\_MYASE.2; UNKNOWN.1.  
 DR PROSITE; PS50125; GUANYLATE\_CYCLASES.2; 1.  
 QO SEQUENCE 690 AA; 78736 MW; 4b736D7413E169AA CRC64;

[illegible]

QY 582 FHLEHGPVSMGKKKEPMQVFLSRKN 608  
: : | : : | |  
Db 571 YKIERGEIDVKGKGTGTWLEEREN 597

Search completed: June 27, 2003, 13:04:43  
Job time : 33.2101 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 27, 2003, 13:03:32 ; Search time 12.7678 Seconds

(without alignments)  
1426.467 Million cell updates/sec

Title: US-09-762-767a-4

Perfect score: 3231  
Sequence: 1 MTGFVNHALLELVIRNTPPE.....QVFLSRKNTGTETKQDD 619Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Issued Patents, AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	404	12.5	1075	1	US-07-623-033-2
2	344.5	10.7	1253	4	US-08-864-785-2
3	309.5	9.6	1134	3	US-08-726-214-2
4	309	9.6	533	3	US-08-726-214-18
5	308.5	9.5	1248	3	US-08-726-214-16
6	291.5	9.0	1294	4	US-09-473-717-2
7	291.5	9.0	1305	4	US-08-864-785-3
8	291.5	9.0	1353	3	US-08-894-173-2
9	291.5	9.0	1353	4	US-09-398-193-2
10	291.5	9.0	1353	4	US-09-473-717-3
11	288	8.9	1104	3	US-08-726-214-10
12	286.5	8.9	1144	3	US-08-726-214-6
13	283.5	8.7	1353	4	US-09-398-193-99
14	282.5	8.7	1099	3	US-08-726-214-14
15	282.5	8.7	1168	4	US-09-474-076-2
16	281	8.7	1064	3	US-08-726-214-8
17	280	8.7	1180	3	US-08-726-214-12
18	278.5	8.6	604	4	US-09-008-097-4
19	278.5	8.6	1167	4	US-09-008-097-6
20	278	8.6	1261	4	US-09-473-716-2
21	272.5	8.4	1165	1	US-08-240-357-2
22	271.5	8.4	1090	3	US-08-307-896-3
23	271.5	8.4	1090	3	US-08-726-214-4
24	271.5	8.4	1090	5	PCR-US95-11808-3
25	269.5	8.3	1077	4	US-09-412-210-1
26	252.5	7.8	429	3	US-08-307-896-4
27	252.5	7.8	429	5	PCR-US95-11808-4

28	175.5	5.4	670	2	US-08-997-080-178	Sequence 178, App
29	175.5	5.4	670	2	US-08-997-362-178	Sequence 178, App
30	175.5	5.4	670	4	US-09-095-855-178	Sequence 178, App
31	175.5	5.4	670	4	US-09-324-542-178	Sequence 178, App
32	175.5	5.4	670	4	US-09-205-426-178	Sequence 178, App
33	175.5	5.4	722	2	US-08-997-080-174	Sequence 174, App
34	175.5	5.4	722	2	US-08-997-362-174	Sequence 174, App
35	175.5	5.4	722	4	US-09-095-855-174	Sequence 174, App
36	175.5	5.4	722	4	US-09-324-542-174	Sequence 174, App
37	175.5	5.4	722	4	US-09-205-426-174	Sequence 174, App
38	173	5.4	419	2	US-08-997-080-125	Sequence 125, App
39	173	5.4	419	2	US-08-997-362-125	Sequence 125, App
40	173	5.4	419	4	US-09-095-855-125	Sequence 125, App
41	173	5.4	419	4	US-09-324-542-125	Sequence 125, App
42	173	5.4	419	4	US-09-205-426-125	Sequence 125, App
43	172.5	5.3	112	4	US-08-857-076-75	Sequence 75, App
44	149	4.6	107	4	US-08-857-076-76	Sequence 76, App
45	137.5	4.3	88	3	US-08-894-173-62	Sequence 62, App

## ALIGNMENTS

RESULT 1  
US-07-623-033-2  
Sequence 2, Application US/07623033  
Patent No. 5237051  
GENERAL INFORMATION:  
APPLICANT: Garbers, David L.  
TITLE OF INVENTION: CLONING THE ENTEROTOXIN RECEPTOR  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESSES:  
ADDRESS: TILTON, FALLON, LUNGUS & CHESTNUT  
STREET: 100 South Wacker Drive - Suite 960  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60606-4002  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/623,033  
FILING DATE: 19901206  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohn, Kenneth I.  
REGISTRATION NUMBER: 30,955  
REFERENCE/DOCKET NUMBER: VU9018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 456-8000  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1075 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-623-033-2  
Query Match 12.5%; Score 404; DB 1; Length 1075;  
Best local Similarity 37.8%; Pred. No. 4.8e-11;  
Matches 102; Conservative 40; Mismatches 98; Indels 30; Gaps 8;  
DB 808 VSLKKEGIVEPELVEVITFSDIVGFTTIC-KYST---PMEVYMLNDIKYSFDQIVD 863

OY		463	SRKNPEYKVEGVGKYMTVSGLP-EPCHIHARSICHLALDMEIAGVOVD--GESVO	518
Dd		864	HHD---YKAVETIGDAIVYASGLPFRNRGNRAHVAIDSKAIDLSPMGTFLEHLPLFPW	920
OY		519	ITGIHTGEVVTVGIGQMPRYCLTGNTVNLTSRTETTGEGKINVSERYTNCILMSPENS	578
Dd		921	IRIGHSGPCAAGVYGICKPRCYCLEGDTVTNTRASMESTGLPLRIHMSSSTAIL--RRT	977
OY		579	DPQHLEHRGPVSMKGKREPMQVWLSKRN	608
Dd		978	DCQPLEYERGETTYLNGRGTEETTYLTGMKD	1007
 RESULT 2 US-08-864-785-2 Sequence 2, Application US/08864785A Patent No. 6329566 GENERAL INFORMATION: APPLICANT: Kaplan, Joshua M. APPLICANT: Oppenheimer, Allison J. APPLICANT: Hart, Anne C. TITLE OF INVENTION: METHODS FOR THE DETECTION, TREATMENT, FILE REFERENCE: 00786/353001 CURRENT APPLICATION NUMBER: US/08/864,785A CURRENT FILING DATE: 1997-05-29 NUMBER OF SEQ ID NOS: 3 SOFTWARE: FastSeq for Windows Version 3.0 SEQ ID NO 2 LENGTH: 1253 TYPE: PRT ORGANISM: Caenorhabditis elegans US-08-864-785-2				
 Query Match            10.7%; Score 344.5; DB 4; Length 1253; Best Local Similarity    32.5%; Pred. No. 5.6e-25; Matches    93; Conservative    45; Mismatches    85; Indels    63; Gaps    8;				
OY		378	LTLRALDEKKKTDLILYSVLPPSVANEL-----	406
Dd		239	LARDLELETPFKDHMIOSVMPKRYADELDKDSLRLRPASNDSCRTSNATQVODPLA	298
OY		407	-----RHRRPVPAKRXYDNTILFSGIVGFNAFCSKHASGEGAMKIYNLNDLYTRFDTL	460
Dd		299	KMPEYRKRFPTFMNLMTNVSLTFADJAGF-----TKMSNNKSADDELVALINDLDFERPDTL	354
OY		461	TDSKRNPVYVEVYVGDKIMYVSGIPCECIHHARISICHIALDAMEIAQOVOD-GESVOI	519
Dd		355	CRLGG--LEKISTLGDCYCVCVAGPEPCDDHACTVMGMDMAIIRQIFDIRQOEYNM	411
OY		520	TIGITHGEVVTVGIGQMPRYCLTGNTVNLTSRTETTGEGKINVSERYTNCILMSPENS	575
Dd		412	RVGHTGTGVMCGMTGRKFKEFDVESNDVTLANEMESSVAGRVHVSSEITAKILKLEYE	471
OY		576	ENSDPQHLEHRGPVSM-----KGKREPMQVWL-SRKNVTGEE	613
Dd		472	EGRP-----YDGFLRMQVQCTERRVRKPKESMKTFPIGRINDGYAE	511
 RESULT 3 US-08-726-214-2 Sequence 2, Application US/08726214 Patent No. 6107076 GENERAL INFORMATION: APPLICANT: Tang, Wei-Jen APPLICANT: Gilman, Alford G. TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE TITLE OF INVENTION: AND USES THEREFOR NUMBER OF SEQUENCES: 31 CORRESPONDENCE ADDRESS: ADDRESSEE: Arnold, White & Durkee STREET: P.O. Box 4433				

CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/726,214  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,498  
FILING DATE: 04-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UTSD:450  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1134 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-726-214-2

Query Match	9.68	Score 309.5	DB 3	Length 1134
Best Local Similarity	33.18	Pred. No. 1.5e-21		
Matches 80	Conservative 49	Mismatches 80	Indels 33	Gaps 8

QY	354	ILGEOREBEXKTLGHELLTDLDTOLTLALDEDEKKPTLLXSVLPSVAHEL	407
Db	235	ILERNORAKFL-QARNCIEDRLR-----LEBENKORILLMSLIPRVAHEKMEDEFLKP	288
QY	408	-----HKRPVPARKYDNTLTLSGIVGFNAFCSKHASEGAMKIVNLINDLYTFPDTLTD	462
Db	289	PERIFPK-----IYIORHDSVILFPADLYGFTGLASQCT-----AQELVKLIMLELFGFDELAT	342
QY	463	SRKNRPVYVNEYGVGDKYMTVSLPRPCIHARSICHLADLME-----IAGVOYDGSSEQ	518
Db	343	EN---HCRNKKILGDCYYVSGSLTQPKDHDHACCVEMGLMDIDTTVSAETVED---LN	396
QY	519	ITIGITHEVYTVIGIORAPRYCLGNTVNLISRTETGEGKATINSEYTRCLMSPEHS	578
Db	397	MRYGLHTGRVLCGVGLRKMQYDVMNSDVTLLANVMEAAAGLRGKVIHTTTLATCLNGSEYV	456
QY	579	DP 580	
Db	457	EP 458	

US-RESULT 4  
US-08-726-214-18  
Sequence 18, Application US/08726214  
Patent No. 6107076  
GENERAL INFORMATION:  
APPLICANT: Tang, Wei-Jen  
APPLICANT: Gilman, Alfred G.  
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENOVIRAL CYCLASE  
TITLE OF INVENTION: AND USES THEREFOR  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/726,214  
 FILING DATE: Concurrently Herewith  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/005,498  
 FILING DATE: 04-OCT-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Highlander, Steven L.  
 REGISTRATION NUMBER: 37,642  
 REFERENCE/DOCKET NUMBER: UTSD:450  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (512) 418-3000  
 TELEFAX: (512) 474-7577  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 533 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-08-726-214-18

Query Match 9.6%; Score 309; DB 3; Length 533;  
 Best Local Similarity 33.6%; Pred. No. 4,6e-22;  
 Matches 77; Conservative 45; Mismatches 75; Indels 32; Gaps 7;

QY 367 QLEFLTRQLTALDEKKKDTLLYSVLPSPVANELR-----HKRPVPAC 415  
 DB 12 QARNCIEDRLR-----LEDENEKORLLMSLPRVAVAMEKEDFLKPERIPHK--IYIQ 64  
 QY 416 RYDNTLIFSGIVGFNAFCSKASGEMAKYVNLNDLYTRPDTLDSKKNFYKVEIV 475  
 DB 65 RDNVSLIFADIVGTGLASQCT-----AOELVKLINELEGFDELATEN--HCRRIKTL 117  
 QY 476 GDKVYVSGLEPFCIHRAISICHLALDME-----IAGOVVDESVOITIGIHGEVYWG 531  
 DB 118 GDCYCVSGSLQPKDHAHCCEMGLMDITISVAENEDV--LNMKVGILHTRVLCG 174  
 QY 532 VIGOMPRYCLEGNTVNLSTRTETGEGKINVSERYTRCLMSPNSDP 560  
 DB 175 VIGLRKQVYVMSNDVTLANVMEAGLPKGVHITKTTLACLNGDYEVER 223

RESULT 5  
 US-08-726-214-16  
 Sequence 16, Application US/08726214  
 Patent No. 6107076  
 GENERAL INFORMATION:  
 APPLICANT: Tang, Wei-Jen  
 APPLICANT: Gilman, Alfred G.  
 TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENITYL CYCLASE  
 NUMBER OF INVENTION: AND USES THEREFOR  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Arnold, White & Durkee  
 STREET: P.O. Box 4433  
 CITY: Houston  
 STATE: Texas  
 COUNTRY: United States of America  
 ZIP: 77210  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/726,214  
 FILING DATE: Concurrently Herewith

CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/005,498  
 FILING DATE: 04-OCT-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Highlander, Steven L.  
 REGISTRATION NUMBER: 37,642  
 REFERENCE/DOCKET NUMBER: UTSD:450  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (512) 418-3000  
 TELEFAX: (512) 474-7577  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1248 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-08-726-214-16

Query Match 9.5%; Score 308.5; DB 3; Length 1248;  
 Best Local Similarity 25.3%; Pred. No. 2,2e-21;  
 Matches 136; Conservative 86; Mismatches 209; Indels 107; Gaps 22;

QY 133 VAOQIH---GTEIDMKVIOQRNECDHT-----QFLIEEKSEKEDFYED 194  
 DB 651 IAQALHVOSGPE-----EINKRIEHTIDRSKDLRREHRIKPSIMKDSLSLEKYSQ 703  
 QY 195 L--DREENGCT-----QESRISPTFCAPFPHIIFDDIVVTOCGNAI 236  
 DB 704 MDEVFYSKLVCAFIYLFETIAIOSLPSRSIMPMTI--QESILMLSHALVITTAED- 760  
 QY 237 YVFLPOLQPGNSLSVSFLVR-----PHIDISFGILSHINTVYLRSEKGLDYEKLE 291  
 DB 761 YKCLPLILKKTCCWINEFYIARVNIIFASILNFGAV--INILMODFQKS--IPLKNLT 816  
 QY 292 CEDELGTETI-----SCLRLKGQITYPEADSLFLCSPSYM-----NLDLITRR 336  
 DB 817 FNSSAVFIDICSYPEYFVETGVLANVTCVFL--RINSVYKLAVLIMIAVALLMETIYA 875  
 QY 337 GYLSDIPLDHNRD-----LVLLGQFPEEYKIQOELITDRQLTLRA----- 382  
 DB 876 GLFLSTDNINHSGEDFLGKREKSLIMAFLLAVFHGOLE-YRARDLFLRWVQAKETI 934  
 QY 383 --LEDEKKKTDFLLYSVLPSPVANELR-----RHKRPVPARKYDNTLIFSGIVGFNAFCSK 436  
 DB 935 NEMKDLREHNENMLRIILFGVABRFLKEDRONNELYSQSYAVGVMAFASIPGPDFYSQ 994  
 QY 437 HASGEGAKIVNLNDLYTRPDTLDSKKNFYKVEIVGDKYVYSGL--PE--PCIHHA 493  
 DB 995 TEMNNOGVCLRLMEIILADFDLGEEDRFODIEKIKTIGSTYMAVSGSPKQOCEDEKW 1054  
 QY 494 RSICHLALDMEIAGOVVDGE---SVQITIGIHGEVYGVIGOMPRYCLFGVTNL 549  
 DB 1055 GHICALADSLALTESIQINKHSFRNFRNFGISGSVYAGVIGAKKFOYIOWGTAVNL 1114  
 QY 550 TSTRTETGEGKINVSERYTRCLMSPNSDPQFHLHRRGVAMKKE---PMQVWFL 604  
 DB 1115 ASRMDSTGVSGRIQVPEERTYLL-----KDGPAFADYREIIVYKISBEGKIKIYFL 1167

RESULT 6  
 US-09-473-717-2  
 Sequence 2, Application US/09473717  
 Patent No. 6372475  
 GENERAL INFORMATION:  
 APPLICANT: Storm, Daniel R.  
 APPLICANT: Hacker, Beth  
 APPLICANT: Tomlinson, James E.  
 APPLICANT: COR Therapeutics, Inc.  
 TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENITYL CYCLASE  
 FILE REFERENCE: 44481-5029-01-US

```

      CURRENT APPLICATION NUMBER: US/09/473,717
      CURRENT FILING DATE: 1999-12-29
      PRIOR APPLICATION NUMBER: PCT/US98/13541
      PRIOR FILING DATE: 1998-07-01
      PRIOR APPLICATION NUMBER: 60/098,559
      PRIOR FILING DATE: 1997-07-01
      PRIOR APPLICATION NUMBER: 08/886,440
      PRIOR FILING DATE: 1997-07-01
      NUMBER OF SEQ ID NOS: 3
      SOFTWARE: PatentIn Ver. 2.0
      SEQ ID NO 2
      LENGTH: 1294
      TYPE: PRT
      ORGANISM: human type IX adenylyl cyclase
      US-09-473-717-2

Query Match          9.0%; Score 291.5; DB 4; Length 1294;
Best Local Similarity 31.2%; Pred.No.12e-19;
Matches   74; Conservative 44; Mismatches 72; Indels 47; Gaps 6

OY    363 LEDEKKKIDTLLYSVLPSPVANEL-----RRK-----              409
DB    323 LEVEALKERKHISVMRPIADLDLMMKGODESESVNKRHATSSPKNRKKKSIOKAPIAF 382
OY    410 RPVPARKRDNTYLIFSGIVGFNAFCSKAHSGBGAKIYNLNDLYTRPDLTDSRKNPV 469
DB    383 RPFMQQIEEVSILFIADIYVF---TKMSANKSAHALVLGLINDLGFRDRLCEETK---C 435
OY    470 YKVGVGGKWMTVSGLPPEPCIHARSICHLALDMMEINGQ-VOVGESVOITIGITHGEV 528
DB    436 EKISTLGACYCVACCPERRADHAACCILEMGIMTKALEDFQEKEKNMAYGHITGY 495
OY    529 VTGVIGORMPRYCFLGNVTNLSTRTTEGCKINVSERYTCMLSPENDSQPHLE 585
DB    496 LCGIIGMRRFFEDWNSDVNIANLMEOLGAVAGKHISEATAKYL-----DDRYEME 546

RESULT 7
US-08-864-785-3
Sequence 3, Application US/08864785A
Patent No. 6329566
GENERAL INFORMATION:
APPLICANT: Kaplan, Joshua M.
APPLICANT: Oppenheimer, Allison J.
APPLICANT: Hart, Anne C.
TITLE OF INVENTION: METHODS FOR THE DETECTION, TREATMENT,
        TITLE OF INVENTION: AND PREVENTION OF NEURODEGENERATION
FILE REFERENCE: 00786/353001
CURRENT APPLICATION NUMBER: US/08/864,785A
CURRENT FILING DATE: 1997-05-29
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 1305
TYPE: PRT
ORGANISM: Mus musculus
US-08-864-785-3

Query Match          9.0%; Score 291.5; DB 4; Length 1305;
Best Local Similarity 31.2%; Pred.No.1.2e-19;
Matches   74; Conservative 44; Mismatches 72; Indels 47; Gaps 6

OY    383 LEDEKKKIDTLLYSVLPSPVANEL-----RRK-----              409
DB    323 LEVEALKERKHISVMRPIADLDLMMKGODESESVNKRHATSSPKNRKKKSIOKAPIAF 382
OY    410 RPVPARKRDNTYLIFSGIVGFNAFCSKAHSGBGAKIYNLNDLYTRPDLTDSRKNPV 469
DB    383 RPFMQQIEEVSILFIADIYVF---TKMSANKSAHALVLGLINDLGFRDRLCEOTK---C 435
OY    470 YKVGVGGKWMTVSGLPPEPCIHARSICHLALDMMEINGQ-VOVGESVOITIGITHGEV 528
DB    436 EKISTLGACYCVACCPERRADHAACCILEMGIMTKALEDFQEKEKNMAYGHITGY 495

```

```

OY      529 VTGVGGQMPRYCGLFGNTVNLTSRFTTGEGKKNVSYYTCCLSPNSDPQHLE 585
       :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      496 LCGILGMRRFDFDWSNDVNLANLMEQLGAVAGKVHISRTAKYL-----DDRYEME 546

RESULT 8
US-08-894-173-2
; Sequence 2, Application US/08894173A
; Patent No. 6090612
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: Adenylylate cyclase and uses therefor
FILE REFERENCE: PL4716C
CURRENT APPLICATION NUMBER: US/08/894,173A
CURRENT FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 97
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 1353
TYPE: PRT
ORGANISM: Mouse
FEATURE:
OTHER INFORMATION: HYPOTHETICAL : NO
US-08-894-173-2

Query Match          9.0%; Score 291.5; DB 3; Length 1353;
Best Local Similarity 31.2%; Pred. No.1,3e-19;
Matches    74; Conservative   44; Mismatches   72; Indels    47; Gaps     6

OY      383 LEDEKKTDLLYSVLPPSVANEL-----RHK-----409
       ||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      323 LEVEKALKERMIHSVMPRILIADDLMKOGDESESNVRKHATSSPKNRKKKSIOKAPIAF 382
OY      410 RPYNAKRKDNTYILFSGLVGNAPCSKIASGBGANIKYNLANDLYTRRDYTLDSKKNFPV 469
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      383 RPFMGOOLEEYSILFDIVGF---TKMSANKSAHALVALGLNLDGRDRDLCOEOTK---C 435
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      470 KYEVTVGDKYMTVSGLPBPCIHARSICHLADMWEINGQ--VVVDGESVOITIGITHGEV 528
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      436 EKISTLGRCYCVAGCCPEPRADHACCEIMGLGMIALKEQCQEKKENVMNRVGHVTGYV 495
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      529 VTGVGGQMPRYCGLFGNTVNLTSRFTTGEGKKNVSYYTCCLSPNSDPQHLE 585
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      496 LCGILGMRRFDFDWSNDVNLANLMEQLGAVAGKVHISRTAKYL-----DDRYEME 546
       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 9
US-09-398-193-2
; Sequence 2, Application US/09398193
; Patent No. 6197581
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: Adenylylate cyclase and uses therefor
FILE REFERENCE: P24360-
CURRENT APPLICATION NUMBER: US/09/398,193
CURRENT FILING DATE: 1999-09-17
NUMBER OF SEQ ID NOS: 104
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 1353
TYPE: PRT
ORGANISM: Mouse
FEATURE:
OTHER INFORMATION: HYPOTHETICAL : NO
US-09-398-193-2

Query Match          9.0%; Score 291.5; DB 4; Length 1353;
Best Local Similarity 31.2%; Pred. No.1,3e-19;
Matches    74; Conservative   44; Mismatches   72; Indels    47; Gaps     6

OY      383 LEDEKKTDLLYSVLPPSVANEL-----RHK-----409
       ||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```



RESULT 12  
US-08-726-214-6  
Sequence 6, Application US/08726214  
Patent No. 6107076  
GENERAL INFORMATION:  
APPLICANT: Tang, Wei-Jen  
APPLICANT: Gilman, Alfred G.  
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE  
TITLE OF INVENTION: AND USES THEREFOR  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/726,214  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/005,498  
FILING DATE: 04-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UTSD:450  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 474-7577  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1144 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-726-214-6  
Query Match 8.9%; Score 286.5; DB 3; Length 1144;  
Best Local Similarity 21.8%; Pred. No. 3.1e-19;  
Matches 119; Conservative 104; Mismatches 201; Indels 121; Gaps 19;  
QY 165 KVIQNEEDCHQFLIEEKSEKEDFEEDLDREENGTOESRISPTCKAPPHIIFD 224  
DB 599 QVKKRNTFLMKRFMDPEMETR---YSVEREKOSGAIFSCSVVLFCTAM-VETLID 652  
QY 225 -----RDVVTOCGNAIYRVLPOLP-----GNCGLS 252  
DB 653 PWIMTYVTFVGEVLLILITICSM--AIFPRAFRKKVAFSSWIDRTMAINTAMLA 710  
QY 253 VESLVRPH-IDT-----SFGILSHINTVFLRSK 281  
DB 711 IFILVANYVDMSCIQYWMGPYNTTIGLDGCMENPKXYNVAVLSIATIMLVQ-- 768  
QY 282 EGLDVEKCEDELTEISCLRLKGOMIYLPDASILFLCSPSV--NDDDLFRG 337  
DB 769 --VSHMKRLTMLLVGA-VTAINLVAMCPVDEYDHRKROEKDSPVALERKQVSTPG 825  
QY 338 LVLSD--IPLDATTDL-VLLGEQFREYKLTQELITLRLQTLALDEKK----- 389  
DB 826 LNTGDSRLPLVPKYSMTVMFMSTFYFSRHVEKLTATLMLKEVLDQEKRYEAM 885  
QY 390 --TDTLISVLPSPVANEL---RHKRPYAKRYDVTILFSGIVGFNAFCRHSAGEA 443  
DB 886 RMEALVTNMLPEHVARHFLGSKRDELSQSYDEIGVAFSLPNRADPTYESINNG 945  
QY 444 MKIVNLNDLITRFDTLDSRKNPFYKVTVDKMTVSGLPDPCIH----- 491

DB 946 IECLRFLEIISPDFSLDMPKFRVITKIKTIGSYMAAGV-TPVNTNGFTSSKEEK 1004  
QY 492 -----HARSICHLDMMEIAGOVVDG-ESVQITIGITHEBVTGCGMPRCLE 543  
DB 1005 SDEKROHLADLDFALMADOTLNTINOSFNNFMILGNKKGVLGAGVIGARRPHDIW 1064  
QY 544 GNTVNLSTRTEGTGKINVSERYRCLMSPNSDPFHLEHGRPVSKRKEPMQWPF 603  
DB 1065 GNTVNVASRMESIGVGNIGVETGYTL-----REYGFVRVGRPIFVAGKE-LITTF 1118  
QY 604 LSRKN 608  
DB 1119 LKGRD 1123  
RESULT 13  
US-09-398-193-99  
Sequence 99, Application US/09398193  
Patent No. 6197581  
GENERAL INFORMATION:  
APPLICANT: Medical Research Council  
TITLE OF INVENTION: Adenylyl cyclase and uses therefor  
FILE REFERENCE: P24360-  
CURRENT APPLICATION NUMBER: US/09/398,193  
CURRENT FILING DATE: 1999-09-17  
NUMBER OF SEQ ID NOS: 104  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 99  
LENGTH: 1353  
TYPE: PRT  
ORGANISM: Human  
US-09-398-193-99  
Query Match 8.8%; Score 283.5; DB 4; Length 1353;  
Best Local Similarity 30.8%; Pred. No. 8.2e-19;  
Matches 73; Conservative 44; Mismatches 73; Indels 47; Gaps 6;  
QY 383 LEDEKKTDTLTVLPSVANEL-----RHK----- 409  
DB 323 LEVKAALKERHSHVMPRIIADLMDKGDESESNVYKRAHTSSPKNRKKSIOKAPIAF 382  
QY 410 RPYPAKRYDNTVTLFSGIVGFNAFCRHSAGEAMKIVNLNDLITRFDTLDSRKNPFY 469  
DB 383 RPKMOOIEVYSILFADIVGF---TRMSANKSAHALVGLNDLFGFRFLCEETK---C 435  
QY 470 KYVEYGDKRYMTVSGLPDPCIHARSICHLDMMEIAGOVVDGSEVQITIGITGEY 528  
DB 436 EKISTGDCYCYACGPEPADAHYCCIEGLMIAIEDFOCKEKEMNRRVGHITRY 495  
QY 529 VTGVIGORMPRCYLFGNTVNLSTRTEGTGKINVSERYRCLMSPNSDPFHLE 585  
DB 496 ICGILGMRBRKFPVWSNDVNLAMLMQDLGAVGVHSEAVAKYL-----DDRYEME 546  
RESULT 14  
US-08-726-214-14  
Sequence 14, Application US/08726214  
Patent No. 6107076  
GENERAL INFORMATION:  
APPLICANT: Tang, Wei-Jen  
APPLICANT: Gilman, Alfred G.  
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE  
TITLE OF INVENTION: AND USES THEREFOR  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/726,214  
 FILING DATE: Concurrently Herewith  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/005,498  
 FILING DATE: 04-OCT-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Highlander, Steven L.  
 REGISTRATION NUMBER: 37,642  
 REFERENCE/DOCKET NUMBER: UTSD:450  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (512) 418-3000  
 TELEFAX: (512) 474-7577  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1099 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-08-726-214-14

Query Match 8.7%; Score 282.5; DB 3; Length 1099;  
 Best Local Similarity 29.3%; Pred. No. 7.2e-19;  
 Matches 84; Conservative 55; Mismatches 97; Indels 51; Gaps 9;

```

QY 334 TRRGILYSDIPLDATRLVLLGEOFFREY-----KITOELEITDRLQTLRLLEDEK 387
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 176 TQLGQL-----LANAVILLGNGFTGAFHKLQDASRLFIYTVACIOIRRLRYEK 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 388 KKTDTLYSVLPSPYANEL-----RHKR-----VPKRYDVTILFSGIV 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 229 ROENLLSVLPRAHSMKALAIIRLKEGDRHMPDNNHSLVYKRNVSILYADIV 288
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 429 GFNAFCSKHAEGEGAMKIVNLNDLYTRFDLTDSRNKPFYKVTVDGKYMTVSGLPEP 488
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 289 GF-----TLASDCSPKELVYVNLFLFKFDIAKANE---CMRIKILGDCYCVSGLPVS 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 489 CIHARSTICHLALDMMETAGOV-QYDGSVOITIGIHTEVVTGYIGQMPRYCLFGNTV 547
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 342 LPTNARNCKVKGGLDICEAIKQVREATGVDISMRVGHSGNYLCGVIGLRKWOYDWMSHDV 401
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 548 NLTSTRTTGGKINVSSEYTYRCIMSPENDPOF-----HLEHRGP 589
   : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 402 SLANMEAGVPGRVHTEATLNL-----DKAYEVEDGHEGRDP 442
   : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 15  
 US-09-474-076-2  
 : Sequence 2, Application US/09474076  
 : Patent No. 6465237  
 : GENERAL INFORMATION:  
 : APPLICANT: Tomlinson, James E.  
 : TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL  
 : TITLE OF INVENTION: CYCLASE  
 : FILE REFERENCE: 44481-5028-01-US  
 : CURRENT APPLICATION NUMBER: US/09/474,076  
 : CURRENT FILING DATE: 1999-12-12  
 : PRIOR APPLICATION NUMBER: PCT/US98/13694  
 : PRIOR FILING DATE: 1998-07-01  
 : PRIOR APPLICATION NUMBER: 60/070,904  
 : PRIOR FILING DATE: 1997-07-01  
 : PRIOR APPLICATION NUMBER: 08/886,550  
 : NUMBER OF SEQ ID NOS: 2  
 : SOFTWARE: Patentin Ver. 2.0  
 : SEQ ID NO 2  
 : LENGTH: 1168

TYPE: PRT  
 ORGANISM: human type VI adenylyl cyclase  
 US-09-474-076-2  
 Query Match 8.7%; Score 282.5; DB 4; Length 1168;  
 Best Local Similarity 26.6%; Pred. No. 8e-19;  
 Matches 94; Conservative 61; Mismatches 126; Indels 73; Gaps 12;

```

QY 296 LVTETISCLRL-----KGMIYLP--ADSIILCSPSVANLMDLTRRGILYSDIPLH 346
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 262 LSGGLSTLHLILAMQLNKRGDAFLMKQGANVLLFCT----- 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 347 DATRLVLLGEOFFREYKLTQLELITDRLQTLRLADEKKKTDLLYSVLPSPYANEL 406
   : : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
DB 300 ---NVIGICHTYPAFVSQRQAFQETRGVQARLH-LQHNROQRLLSLVLPQVAMEM 354
   : : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
QY 407 R-----HRRPYARKYDVTILFSGIVGFNPFCSKHAEGEGAMKIVNLNDLY 454
   : : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
DB 355 KEDITKKEDMMFHK--IYIQKHNVSTILFDIEGFTSLASQCT---AQELVMTLNELF 408
   : : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
QY 455 TRFDLTDSRNKPNFYKVTVDGKYMTVSGLPEPCIHARSICHLALDMMETAGOV-QYD 513
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 409 ARFDKLAEN---HCLRIKILGDCYCVSGLPEARADAHCCVEKGVDMIEAISLYREY 465
   : : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
QY 514 GESVOITIGIHTEVVTGYIGQMPRYCLFGNTVNLTSRTTETGEGKINVSSEYTYRCIM 573
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 466 GVNVMNRVGIHSGRVHCGVIGLRKWOYDWMNDVTLANHMEAGRAGRILHTRATLQYLN 525
   : : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
QY 574 SPENDP-----QFHLH---RGVPNKKKKKKEMQVFLSRKNTGEE 613
   : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 536 GDYEVEPGRGERNAYLKEOHITETLLIG-ASQKRREKAMIAKLQRRRANSE 578
   : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: June 27, 2003, 13:05:59  
 Job time : 14.7678 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 13:05:32 ; Search time 32.6287 Seconds

(without alignments)  
2080.112 Million cell updates/sec

Title: US-09-762-767a-4

Perfect score: 3231  
Sequence: 1 MYGFVNHAELELVIRNTPPE.....QVFLSRKNTGTETKQDD 619

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 10964683 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCY\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCYS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	789.5	24.4	690	9	US-10-205-823-162
2	778	24.1	691	9	US-09-952-213D-2
3	407.5	12.6	1073	10	US-09-819-249-2
4	405.5	12.6	1073	9	US-10-157-031-18
5	291.5	9.0	1294	12	US-10-071-223-2
6	291.5	9.0	1353	10	US-09-751-100B-2
7	291.5	9.0	1353	12	US-10-071-223-3
8	289	8.9	1080	9	US-10-282-942-2
9	283.5	8.8	1353	10	US-09-751-100B-99
10	282.5	8.7	1168	9	US-10-201-000-2
11	282.5	8.7	1168	10	US-09-750-240-11
12	278.5	8.6	604	10	US-09-750-240-4
13	278.5	8.6	1167	10	US-09-750-240-6
14	278	8.6	1261	9	US-10-175-136-2
15	269.5	8.3	257	9	US-09-989-442-95
16	269.5	8.3	302	10	US-09-915-582-50
17	265.5	8.3	1077	10	US-10-121-911-1
18	265.5	8.2	1167	10	US-09-750-240-13
19	264	8.2	241	9	US-09-764-868-794

20	264	8.2	241	9	US-09-989-442-141	Sequence 141, App
21	259.5	8.0	855	10	US-09-925-297-911	Sequence 811, App
22	242	7.5	225	9	US-09-989-442-92	Sequence 92, App1
23	234.5	7.3	311	10	US-09-925-297-515	Sequence 515, App
24	175.5	5.4	670	9	US-10-051-643-178	Sequence 178, App
25	175.5	5.4	670	9	US-09-880-505-178	Sequence 178, App
26	175.5	5.4	722	9	US-10-051-643-174	Sequence 174, App
27	175.5	5.4	722	9	US-09-880-505-174	Sequence 174, App
28	173	5.4	419	9	US-10-051-643-125	Sequence 125, App
29	173	5.4	419	9	US-09-880-505-125	Sequence 125, App
30	172.5	5.3	112	10	US-09-205-658-75	Sequence 75, App1
31	172.5	5.3	112	10	US-09-844-353A-75	Sequence 75, App1
32	149	4.6	107	10	US-09-205-658-76	Sequence 76, App1
33	149	4.6	107	10	US-09-844-353A-76	Sequence 76, App1
34	137.5	4.3	88	10	US-09-751-100B-62	Sequence 62, App1
35	137	4.2	85	10	US-09-751-100B-59	Sequence 59, App1
36	137	4.2	85	10	US-09-751-100B-60	Sequence 60, App1
37	137	4.2	85	10	US-09-751-100B-61	Sequence 61, App1
38	134.5	4.2	90	10	US-09-751-100B-49	Sequence 49, App1
39	131.5	4.1	90	10	US-09-751-100B-48	Sequence 48, App1
40	131	4.1	24	9	US-10-106-698-5238	Sequence 5238, App1
41	131	4.1	85	10	US-09-751-100B-54	Sequence 54, App1
42	131	4.1	85	10	US-09-751-100B-55	Sequence 55, App1
43	131	4.1	85	10	US-09-751-100B-56	Sequence 56, App1
44	131	4.1	85	10	US-09-751-100B-57	Sequence 57, App1
45	131	4.1	85	10	US-09-751-100B-58	Sequence 58, App1

#### ALIGNMENTS

RESULT 1  
US-10-205-823-162  
; Sequence 162, Application US/10205823  
; Publication No. US20030108963A1  
GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Ganavarapu, Manjula  
; APPLICANT: Goratcheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamalkar, Shubhangi  
; APPLICANT: Monsey, Angela M.  
; APPLICANT: Glat, Karen  
; APPLICANT: Zhao, Xumel  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
THERAPY OF PROSTATE CANCER  
FILE REFERENCE: MRI-044  
CURRENT FILING DATE: 2002-07-25  
PRIOR APPLICATION NUMBER: US/10/205, 823  
PRIOR FILING DATE: 2001-07-25  
PRIOR APPLICATION NUMBER: 60/314, 356  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/325, 020  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: 60/341, 746  
PRIOR FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: 60/362, 158  
PRIOR FILING DATE: 2002-03-05  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 162  
LENGTH: 690  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-205-823-162  
Query Match 24.4%; Score 789.5; DB 9; Length 690;  
Best Local Similarity 35.1%; Pred. No. 5.3e-60;

Matches	207: Conservative	96: Mismatches	219: Indels	67: Gaps	15:
Oy	30	OLDEGCGATVRIYDDSTYDLVAASVYLNINAGEIL-QMGFKMEFFPCGSGSDITLIR	88		
Dd	101	KIKESRKELEHDEFKTTIAEQVAVGV-----EVIKESIGSEVFFICYED-ENTLLG	153		
Oy	89	VLSNVREELONLDAL-----HDILATITPGMKRAPSFCCTAERKGGILHYYSREBGLQ	143		
Dd	154	VVGGLIKDFLNSFSTLLQSSHCQAGKRGRLDASITCLDKE-DDFLHYVYFFPKRTTS	212		
Oy	144	DIVIGITVAAQIHGTETIDMKVIOQRNEEDCHFO-----FLIEKESKEEDPFY	192		
Dd	213	LILPILITAAAHVITYETEVEVSLM---PPCFHNDSCSEFVNOPYLLYSVNHKSTKPS---	265		
Oy	193	EDLDPEENGTOESRISPYT-FCKAFPHIIFDRDLVVTGCGNATVRIYLPOLQ-PGNCSL	250		
Dd	266	-----LSPSKQSSSLVITPSTLECKTFPHFPHFDKMDITLQNGNGLRLMNRDPOGAKNF	320		
Oy	251	LSVFSVLVRPHIDISFHGILSHINTVFVLSRKEGLLDVKELECEDELGTETISCLRLKGQM	310		
Dd	321	EVEYFELLPKINQFSGIMTLMNQFVYVRVRMDNSVAK-----SSRVMDLKGQM	370		
Oy	311	IYLPADSIILFLCSPSVNNLDDLTRRGVLYSLDIPHDTRDLYLLGEOPREBYKLTGLE	370		
Dd	371	IYVESSALILPLGSCVDRLDEFGRGLYSLDIPHNLRVYLLGEOPARQDGLKRLIG	430		
Oy	371	ILTRRLQALTRALDEBKRRKTDITLYSVLRPPSVAMELRHRKRVPAKRYDNVITLFSGIVGF	430		
Dd	431	KLKATLEQAOHALEREKKRYDLCISIPCEVAQQLMGQVQAOKFSSVMTMLPFSIDYGF	490		
Oy	431	NAPCSKHAASGCAKIVYLLNDLYTRFPTLYDSKRNPFVYAYEYVGDYKITYVSGLPFECI	490		
Dd	491	TALCSQCS-----PLQVITMLNALTYTRFOCGEILD---VYVETTGDAYCVAGGLHKESD	543		
Oy	491	HHARSICHALDMMELIAGQV-OVDGESVQITIGIHTGSEVTVGIVIGORAPRYCLEFENTVL	549		
Dd	544	THAVALMALKMDLSEVMSPHEP LKMRIGLHSGSVFVFGVYKMRPRCLCENNNTL	603		
Oy	550	TSRTEITTEKKGKINVSETYRCL-----MSPENDSQPHLEHRNG	588		
Dd	604	ANKFSCSVPRKINVSPTTYRLKLDCCPGFVTPPRSRELPNPFSEIPG	652		
RESULT 2					
US-09-952-213D--2					
Sequence 2, Application US/09952213D					
Publication No. US20030096240A1					
GENERAL INFORMATION:					
APPLICANT: MURAD, FERID					
APPLICANT: SHARINA, IRAIDA G.					
APPLICANT: KRUMENACKER, J. S.					
APPLICANT: MARTIN, E.					
TITLE OF INVENTION: GENOMIC ORGANIZATION OF MOUSE AND HUMAN SGC					
FILE REFERENCE: UTSH:2520S					
CURRENT APPLICATION NUMBER: US/09/952.213D					
CURRENT FILING DATE: 2002-08-16					
NUMBER OF SEQ ID NOS: 15					
SOFTWARE: PatentIn Ver. 2.1					
SEQ ID NO 2					
LENGTH: 691					
TYPE: PRT					
ORGANISM: Mus musculus					
US-09-952-213D-2					
Query Match					
Best Local Similarity 33.5%, Pred. No. 5.3e-59;					
Matches 213: Conservative 105; Mismatches 224; Indels 94; Gaps 20;					
Oy	9	LELVIRNYGCEGVEDIKKEAQLDEGQFLRIIYDDSTYDLVAASKVYLNINAGEITLQ	68		
Dd	89	LNIALQRIALAKHKIEENKRSSEKED-----LEKIIAER-----IAAGAPVPAALK-----D	134		
Oy	69	MFGMAFFVCOESGSDITLIRLVGNSVREELONLDAL-----HDILATITPGMKRAPSFCCT	123		

```

Db      135 SLGELFKICYED - EHLGVGVGTLKDLNFSPTLLKSSCHQEAERGRLEDASTL 193
QY      124 DAERKGLILHYSSREGLDIVIGIKTVAQOIHGTEIDMKVIOQRNEC---DHTOPL 180
Db      194 KDKQ-DENVVYFFFKRTTALLPGIKMAAILVESHVYSIM---PPCFSDCTERY 248
QY      181 -----IEEKSEDEYEDLDNRENGTQESRISPT-PCNAFFHIIIDROLVYMO 231
Db      249 NOPILLYSVHAKSTPS-----LSPEKPOSSLVTPSLCKTFPFHFMIDROLATIQ 300
QY      232 CGMALRYVLPOLQ-EGNCSLSVFSLYRPHIDISFGILSHINTVYFLRSKRGILDVEXT 290
Db      301 LGNGIRRLVYNRKDFQKPRFEFEFFELLPKINOTFSGIMTMNGVIVRYBRMDSYVK- 359
QY      291 ECEDDLTGTETISCLAKGMYILPNDLSILFCSPSYMHLDLTRGLGLSDIPLDADR 350
Db      360 -----SSRYMOLKGOMIYIVESSALLFSPVDLEDEFTGGLYLSDIPIHNAKR 410
QY      351 DVLVLGEQFREERYKLTQLELEITDRLQTLRLALEDEKRTDPLVSLVLPSPANELRHR 410
Db      411 DVLVLGEQBARADGLKRLRGKATLEHNAQLAEBEKRTVOLLCSIFPSYVAQOIMOG 470
QY      411 PVPARKYDNTVILFSGIVGFENAFCSKHASEGAMKIVNLNDLYTRFDTLDSRKNPEY 470
Db      471 IVOAKKFSEVITLFSDIYGFATICSOS---PLAYITMLNLYTRFFDOCGELD--Y 523
QY      471 KVEYVGDDYMYTVSGLPBPCITHARSLCHALDMMETAGOV-QVDSGVSVOITIGHGEV 529
Db      524 KVEITGDAYCVAGGLHRESDTHAVOIALMALKMLESNEMVSPHGPPIKRIGLHSGSYF 583
QY      530 TGVIGMRPRYCLFSGNTVNLISRTETTGKRGKINVEYTRCL-----MSPEND 579
Db      584 AGVGVKMPRYCLFSGNTVNLANKFESCSVPRKINVSPTTYRLKDCPGVFTPRSRELP 643
QY      580 POF-----HL-----EHRGPVSMKKGKEPMOYF 603
Db      644 PNFPSDIPGICHTFLDAYHHQPN-----KPMF 671

RESULT 3
US-09-819-249-2
Sequence 2, Application US/09819249
Patent No. US20010029019A1
GENERAL INFORMATION:
APPLICANT: Waldman, Scott A.
APPLICANT: Park, Jason
APPLICANT: Schulz, Stephanie
TITLE OF INVENTION: Compositions And Methods For Identifying And Targeting Cancer
FILE REFERENCE: TUD2412
CURRENT APPLICATION NUMBER: US/09/819, 249
CURRENT FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: 60/192, 229
PRIOR FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1073
TYPE: PRT
ORGANISM: Homo sapiens
US-09-819-249-2

Query Match      12.6%; Score 407.5; DB 10; Length 1073;
Best Local Similarity 22.6%; Pred. No. 2e-26; Indels 225; Gaps 26;
Matches 180; Conservative 117; Mismatches 274;

12 LVIRNYGEEVEDIKREAQLDEGQFLVRIYIDSDSKTYDLVAAS-----KYLNLNAGEI 66
::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::| 111
Db 245 VIMOGGSEFLYKLKGDAVAEDYIILVDLFNDQYLEDNVTAPPYMKRVLYLITLSPGS 304
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::| 111
QY 67 L-----OMFGKMFVFQCF-----SGYD 84
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::| 111

```

```

Db 305 LANSFSRNLSTPKDFRLAYINGLIYVGHMALKIFLNGENITTPKFAHFRNLPEGYD 364
Qy 85 -----TLIRVLSGVNREFLQMDALHDLIAYDGMRAFSRCDAE----- 126
Db 365 GPVTLDMQDVSTVWLYLTSTYDKKRYKLYLTFDFHVNKTVPDMSPTFWKNSKLPNDI 424
Qy 127 KKG---LIIHYSEBEGQDIYIGIKVVAQOIHSTEDMKVIOQR----- 170
Db 425 TGRGQIIMIAVFT---LTGAVALLLVALLMLAKRYRDYELRQKWSHIPENIFPLE 480
Qy 171 NECDHTOFLIE-----KSEKEDFEYDLDRFENGQOESR-----IS 209
Db 481 TMTNHVSLKIDDKRDTIORLROCKYKRYLLKDLKNDGNTFKQKIELNKLQID 540
Qy 210 PYTCKAPFPHIIFRDVLVTCGNAIYVLPOLQPGNSLLSVSLVRPHIDISPHGLI 269
Db 541 YTLTKFYG-----TVKIDPMIFGVIECERG--SLREVLDNTISPGCTMDWE 588
Qy 270 SHINTVYLRKSGGLDVKECEDELTEISCLRLKQMTYLP--ADSL----- 320
Db 589 FRTSVLYDIAGKMSYLAHSSKTEVHGRLKST--NCVDSRMVYKIDFGCNSILPPKDL 646
Qy 321 ----FLCSPVANLDDLTRRGLYLSIDIPLDAT-----RD----- 351
Db 647 TAPFHLRQANISOKGDVSYGIIAQEIIIRKETFYTLSCDRNEKIFRVENSGMRPRP 706
Qy 352 ---LVLLGBOFREKKLQOE-----RD-----LEILD 374
Db 707 DLFLTAERKELEVLIVKNCWEDEKRPDKIETTLAKIFGLPHDQKESYMDTLIR 766
Qy 375 RQULRLALED-----EKKTDLLYSVLRPSYANELRHRPPARKYDVTLL 423
Db 767 RLQLSRNLHEHYEERTOLYKAERADRLNMLPLRLVVKSLKEGFEPELYEVEYIY 826
Qy 424 FSGIYGFNAFCSKHSAGAMIVNLDLYTRFDLTJTSRKNPEYKVTGDKMTYS 483
Db 827 FSDIYGFITIC-KYST---PMEVMDLNDIYKSFHDIVDHD---YKKEITJSDAYWAS 879
Qy 484 GUPF-PCIHARSICHLADMMELIAGQYVD--GESVOITIGINTGEVTVGIGORMR 539
Db 880 GLPKRNGNHAIDIAKMALEILISFMGTFELEHLPGLPIMIRIGVHSGPCAGVYIKMR 939
Qy 540 YLFGNTVNLSTRTETGKGIINSEYTYRCLMSFENDPOFHLEHGRPVSKKKEPM 599
Db 940 YLFGDVTNTASRMESTGLPLRIHVSSTIALI---KRTCOFLYEVREGETYLKGRNET 996
Qy 600 QVWFLSRKNVTGEERK 615
Db 997 TYWL-----TGKDKK 1007

```

```

RESULT 4
US-10-157-031-18
; Sequence 18, Application US/10157031
; Publication No. US20030108690A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashov, A. V.
; APPLICANT: Krutovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US-10/157, 031
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-18

```

```

Query Match 12.6%; Score 405.5; DB 9; Length 1073;
Best Local Similarity 37.2%; Pred. No. 3e-26;
Matches 103; Conservative 44; Mismatches 95; Indels 35; Gaps 9;
Qy 354 LIGGEFREYKLTQLEILLDRQLRLALED-----EKKTDLLYSVLRPSY 402
Db 751 LEHDQKNESY---MDTLIRQLQLSRNLEHYEERTOLYKAERADRLNMLPLRYV 805
Qy 403 ANELRRKRRVPARKYDNTITLFGIYGFNAFCSKHSAGAMIVNLDLYTRFDLT 462
Db 806 VSLKRGKGYEPELYEVEYIYPSDIYGFITIC-KYST---PMEVMDLNDIYKSFHD 861
Qy 463 SRNPFYVYEVYDGRYTVSGIPE-PCIHARSICHLADMMELIAGQYVD--GESVQ 518
Db 862 HHD---VYKVEITGDAYWASGLPKRNGNHAIDIAKMALEILISFMGTFELEHL 918
Qy 519 ITTIGTGEVYVGIQORHPRYGLPONTYNTLSRTETGKGIINSEYTYRCLMSFEN 578
Db 919 IRIGVHSGPCAGVYIKMRPYGLFGDVTNTASRMESTGLPLRIHVSSTIALI---KRT 975
Qy 579 DPOFLEHHRGPVSMKGRKKEPMQVWFLSRKNVTGEERK 615
Db 976 ECOFLYEVREGETYLKGRNETTYWL-----TGKDKK 1007

```

```

RESULT 5
US-10-071-223-2
; Sequence 2, Application US/10071223
; Patent No. US20020137174A1
; GENERAL INFORMATION:
; APPLICANT: Storm, Daniel R.
; APPLICANT: Hacker, Beth
; APPLICANT: Tomlinson, James E.
; APPLICANT: Cor Therapeutics, Inc.
; APPLICANT: University of Washington
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
; FILE REFERENCE: 44481-5029-02-US
; CURRENT APPLICATION NUMBER: US/10/071, 223
; PRIOR APPLICATION NUMBER: 09/473, 717
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PCT/US98/13541
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/098, 559
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 08/886, 440
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1294
; TYPE: PRT
; ORGANISM: human type IX adenylyl cyclase
US-10-071-223-2

```

```

Query Match 9.0%; Score 291.5; DB 12; Length 1294;
Best Local Similarity 31.2%; Pred. No. 3.5e-16;
Matches 74; Conservative 44; Mismatches 72; Indels 47; Gaps 6;
Qy 383 LEDEKKTTDLTVSLRPPSYANL-----RHK----- 409
Db 323 LEVEKALKERMIHSHVPRITADLMDKOGDESESNVYRHAFTSPKRRKKSIOKAP 382
Qy 410 RPYPAKRYDNTITLFGIYGFNAFCSKHSAGAMIVNLDLYTRFDLTJTSRKNPFV 469
Db 383 RPKMQQIEVSLTFLFDIVGE---TKMSANKSAHLVGLLDFGRFRLCEETK---C 435
Qy 470 YKVEYDGMKMTYVSGLPICIHARSICHLADMMELIAGQYVD--GESVOITIGINTGEV 528
Db 436 EKISTIGDCYCYCAAGPEPRADHAACYCIENGLMIRAIHQFOCKEMVMKRGVHTGIV 495
Qy 529 VTGVIQORMPRYGLFGNTVNLISRTETGKGIINSEYTYRCLMSFENDPOFHLE 585

```

Db 496 LCGILGMRKRRKFDVMSNDVNLNLMEDLGAVAKVHISEATAKYL-----DDRYEME 546

RESULT 6

US-09-751-1008-2

Sequence 2, Application US/097511008

Patent No. US20020142436A1

GENERAL INFORMATION:

APPLICANT: Medical Research Council

TITLE OF INVENTION: Human Adenylate Cyclase and Use Therefor

FILE REFERENCE: P27948A

CURRENT APPLICATION NUMBER: US/09/751,1008

CURRENT FILING DATE: 2000-12-28

NUMBER OF SEQ ID NOS: 104

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 1353

TYPE: PRT

ORGANISM: Mouse

US-09-751-1008-2

Query Match 9.0%; Score 291.5; DB 10; Length 1353;

Best Local Similarity 31.2%; Pred. No. 3.7e-16;

Matches 74; Conservative 44; Mismatches 72; Indels 47; Gaps 6;

QY 383 LEDERKKTDTLLYSVLPPSVANEL-----RHK----- 409

Db 323 LEVERKALKERIMHIVMPRIIADLMKQDESENSVSRHATSSPKNRKKSSIOKAPIAF 382

QY 410 RPVAKRYDNTVILFSGIVGNACSKHASEGAMKIVNLDYTRPDTLTDKRKNPFV 469

Db 383 RPFMOQIEVSIILFAIVGF-----TKMSANKSAHALVGLNDLFGFRDLCEQTK---C 435

QY 470 YKVEVDKRYMTVSGLEPDCIHARSICHLALDMMETIAGO-VQYDGESVOITIGIHTEV 528

Db 436 EKISTLDCCYCVAGCEPRADHAYCCIEMLGMIKALIEQCEQKEKEMVNRVGHITGV 495

QY 529 VTGVIGQMRPRYCLFGVTNLTSTETTGEGKINVSSEYTYRCIAMSPEISDPFHLE 585

Db 496 LCGILGMRKRRKFDVMSNDVNLNLMEDLGAVAKVHISEATAKYL-----DDRYEME 546

RESULT 7

US-10-071-223-3

Sequence 3, Application US/10071223

Patent No. US20020137174A1

GENERAL INFORMATION:

APPLICANT: Storm, Daniel R.

APPLICANT: Hacker, Beth

APPLICANT: Tomlinson, James E.

APPLICANT: COR Therapeutics, Inc.

TITLE OF INVENTION: Cloning and Characterization of a Human Adenylate Cyclase

FILE REFERENCE: 44481-5029-02-US

CURRENT APPLICATION NUMBER: US/10/071,223

CURRENT FILING DATE: 2002-02-11

PRIOR APPLICATION NUMBER: 09/473,717

PRIOR FILING DATE: 1999-12-29

PRIOR APPLICATION NUMBER: PCT/US98/13541

PRIOR FILING DATE: 1998-07-01

PRIOR APPLICATION NUMBER: 60/098,559

PRIOR FILING DATE: 1997-07-01

PRIOR APPLICATION NUMBER: 08/886,440

PRIOR FILING DATE: 1997-07-01

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 1353

TYPE: PRT

ORGANISM: murine type IX adenyl cyclase

US-10-071-223-3

Query Match 9.0%; Score 291.5; DB 12; Length 1353;

Best Local Similarity 31.2%; Pred. No. 3.7e-16;

Matches 74; Conservative 44; Mismatches 72; Indels 47; Gaps 6;

QY 383 LEDERKKTDTLLYSVLPPSVANEL-----RHK----- 409

Db 323 LEVERKALKERIMHIVMPRIIADLMKQDESENSVSRHATSSPKNRKKSSIOKAPIAF 382

QY 410 RPVAKRYDNTVILFSGIVGNACSKHASEGAMKIVNLDYTRPDTLTDKRKNPFV 469

Db 383 RPFMOQIEVSIILFAIVGF-----TKMSANKSAHALVGLNDLFGFRDLCEQTK---C 435

QY 470 YKVEVDKRYMTVSGLEPDCIHARSICHLALDMMETIAGO-VQYDGESVOITIGIHTEV 528

Db 436 EKISTLDCCYCVAGCEPRADHAYCCIEMLGMIKALIEQCEQKEKEMVNRVGHITGV 495

QY 529 VTGVIGQMRPRYCLFGVTNLTSTETTGEGKINVSSEYTYRCIAMSPEISDPFHLE 585

Db 496 LCGILGMRKRRKFDVMSNDVNLNLMEDLGAVAKVHISEATAKYL-----DDRYEME 546

RESULT 8

US-10-282-942-2

Sequence 2, Application US/10282942

Patent No. US20030087295A1

GENERAL INFORMATION:

APPLICANT: Sllas-Santiago, Immaculada

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND DIAGNOSIS OF PAIN DISORDERS USING 9805

FILE REFERENCE: MPT01-240PIRA

CURRENT APPLICATION NUMBER: US/10/282,942

CURRENT FILING DATE: 2002-10-29

PRIOR APPLICATION NUMBER: 60/335,047

PRIOR FILING DATE: 2001-10-31

NUMBER OF SEQ ID NOS: 2

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 1080

TYPE: PRT

ORGANISM: Homo sapien

US-10-282-942-2

Query Match 8.9%; Score 289; DB 9; Length 1080;

Best Local Similarity 26.3%; Pred. No. 4.4e-16;

Matches 94; Conservative 66; Mismatches 109; Indels 88; Gaps 12;

QY 236 IYRYLPQLOGNCISLVSFLVRPHIDISFHGILSHINTFYLRSKRGGLDVERKEDE 295

Db 136 VYTLPLPSMKGAVAV-----GAVSTASHLLVLSLNG----- 167

QY 296 LTGTEISCLRKGOMIYLPKADSLFLCSPSVNLDLTRGGLYSDIPHDATRDLYL 355

Db 168 --GFTTSPVVRGLQL--ANAVIFLCG--NLGAFKHK-----QMGDSRDL-- 208

QY 356 GEQPREYKLTQELIITLRLQTLRLADEKKTDTLLYSVLPSSVA-----N 404

Db 209 -----FTYVAKCIOIR-----RKIRIEKROEMILLVLPAAHISMCKALIERLK 254

QY 405 ELRHRKRPV-----AKRYDNTVILFSGIVGNACSKHASEGAMKIVNLDYTR 456

Db 255 EHGDRKMPDNNFHSILYKRNHONSLIYADIVGTOLASDQSPSE-----LVVYLNELFGK 310

QY 457 FDTLTDKRKNPFVYKVEVDKRYMTVSGLEPDCIHARSICHLALDMMETIAGO-VQYDG 515

Db 311 FDOIAKANE---CHRITLDCYCVAGSLVSPETHARNCYKMGDLMOCAIKQVREATGV 367

QY 516 SVQITIGIHTEVTVGVIGQMRPRYCLFGVTNLTSTETTGEGKINVSSEYTYRCL 572

Db 368 DINNRVGHISGNVLCVIGLRKQYDVMSHDVSLANMEAGAVGVRHITTEATLKLH 424

RESULT 9

[illegible]

Db 409 ARPKLAEN---HCLRIKILGDCYVCVSGLPEARADHACVGMVDMIAISLVREVT 465  
QY 514 GESVOITIGHTGEVTVGIGQMRPRYCLGNTVNLTSRRTTGEKKNINSEYTYCLM 573  
Db 466 GVNMMVGHSGHGVGIGLGRKMQFDVMSNDVTLANHMEAGSRAGRIRHITRATLYIN 525  
QY 574 SPENSDP-----QHLEH---KGVYSMAKKEPQWFLSKNKTGTEE 613  
Db 526 GDYEVEERGGERMAYLKEQHIEFTLLIG-ASOKRKEKMLAKIQRTRANSME 578

## RESULT 12

US-09-750-240-4  
; Sequence 4, Application US/09750240  
; Patent No. US20020103147A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammon, H. K.  
; APPLICANT: Insel, P. A.  
; APPLICANT: Ping, P. A.  
; APPLICANT: Post, S. R.  
; APPLICANT: Gao, M.  
; TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART  
; FILE REFERENCE: 220002056723  
; CURRENT APPLICATION NUMBER: US/09/750,240  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 09/472,667  
; PRIOR FILING DATE: 1999-12-27  
; PRIOR APPLICATION NUMBER: US 09/008,097  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: US 08/924,757  
; PRIOR FILING DATE: 1997-09-05  
; PRIOR APPLICATION NUMBER: US 60/048,933  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: US 08/708,661  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 604  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-750-240-4

Query Match 8.6%; Score 278.5; DB 10; Length 604;  
Best Local Similarity 27.1%; Pred. No. 1.5e-15;  
Matches 83; Conservative 56; Mismatches 110; Indels 57; Gaps 9;

QY 296 LQTEISCLRL-----KGOMIYLP--ADSLIFCSPSVNMLDLTRGGLYSDIPLH 346  
Db 51 LSLGLSTHLILAMOLNRGDAPFLMKOLGANVILFCT----- 88  
QY 347 DATRDVLAGEQREKYLQELIIDLRLQLRALDEKRRKTDLLSYLPPSVANEL 406  
Db 89 ---NWSICTHYPALVSQROAFQETRSYIQAHLH-LQHNROQERLLSYLPQHAMEX 143  
QY 407 R-----HKRPYPAKRYDNVTILFSGIVGFNAFCSKHAHSGGAKMIVNLNDLYT 455  
Db 144 KEDINRKEDMFHK--IYIOKHNVSLIFADIGFTSLASQCT---AQELVMTLNEFLA 197  
QY 456 RFDTLDSRKNRPYVEVTVGDKYMTVSGLPEPCIHARSICHLADMMELIAGV-QVDS 514  
Db 198 RPKLAEN---HCLRIKILGDCYVCVSGLPEARADHACVGMVDMIAISLVREVTG 254  
QY 515 ESVOITIGHTGEVTVGIGQMRPRYCLGNTVNLTSRRTTGEKKNINSEYTYCLM 574  
Db 255 VNNMMVGHSGHGVGIGLGRKMQFDVMSNDVTLANHMEAGSRAGRIRHITRATLYIN 525  
QY 575 SPENSDP 580  
Db 315 DYEVEP 320

## RESULT 13

US-09-750-240-6  
; Sequence 6, Application US/09750240  
; Patent No. US20020103147A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammon, H. K.  
; APPLICANT: Insel, P. A.  
; APPLICANT: Ping, P. A.  
; APPLICANT: Post, S. R.  
; APPLICANT: Gao, M.  
; TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART  
; FILE REFERENCE: 220002056723  
; CURRENT APPLICATION NUMBER: US/09/750,240  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 09/472,667  
; PRIOR FILING DATE: 1999-12-27  
; PRIOR APPLICATION NUMBER: US 09/008,097  
; PRIOR FILING DATE: 1998-01-16  
; PRIOR APPLICATION NUMBER: US 08/924,757  
; PRIOR FILING DATE: 1997-09-05  
; PRIOR APPLICATION NUMBER: US 60/048,933  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: US 08/708,661  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1167  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-750-240-6

Query Match 8.6%; Score 278.5; DB 10; Length 1167;  
Best Local Similarity 27.1%; Pred. No. 4e-15;  
Matches 83; Conservative 56; Mismatches 110; Indels 57; Gaps 9;

QY 296 LQTEISCLRL-----KGOMIYLP--ADSLIFCSPSVNMLDLTRGGLYSDIPLH 346  
Db 262 LSLGLSTHLILAMOLNRGDAPFLMKOLGANVILFCT----- 299  
QY 347 DATRDVLAGEQREKYLQELIIDLRLQLRALDEKRRKTDLLSYLPPSVANEL 406  
Db 300 ---NWSICTHYPALVSQROAFQETRSYIQAHLH-LQHNROQERLLSYLPQHAMEX 354  
QY 407 R-----HKRPYPAKRYDNVTILFSGIVGFNAFCSKHAHSGGAKMIVNLNDLYT 455  
Db 355 KEDINRKEDMFHK--IYIOKHNVSLIFADIGFTSLASQCT---AQELVMTLNEFLA 408  
QY 456 RFDTLDSRKNRPYVEVTVGDKYMTVSGLPEPCIHARSICHLADMMELIAGV-QVDS 514  
Db 409 RPKLAEN---HCLRIKILGDCYVCVSGLPEARADHACVGMVDMIAISLVREVTG 465  
QY 515 ESVOITIGHTGEVTVGIGQMRPRYCLGNTVNLTSRRTTGEKKNINSEYTYCLM 574  
Db 466 VNNMMVGHSGHGVGIGLGRKMQFDVMSNDVTLANHMEAGSRAGRIRHITRATLYIN 525  
QY 575 SPENSDP 580  
Db 526 DYEVEP 531

## RESULT 14

US-10-175-158-2  
; Sequence 2, Application US/10175158  
; Publication No. US20030008371A1  
; GENERAL INFORMATION:  
; APPLICANT: Tomlinson, James  
; APPLICANT: Cor Therapeutics, Inc.  
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL  
; FILE REFERENCE: 44481-5027-01-US  
; CURRENT APPLICATION NUMBER: US/10/175,158

PRIOR FILING DATE: 2000-08-22	PRIOR APPLICATION NUMBER: 60/2215, 6647
PRIOR FILING DATE: 2000-07-07	PRIOR APPLICATION NUMBER: 60/2225, 2675
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/2216, 8805
PRIOR FILING DATE: 2000-07-07	PRIOR APPLICATION NUMBER: 60/2225, 2707
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/2251, 8653
PRIOR FILING DATE: 2000-12-08	PRIOR APPLICATION NUMBER: 60/2235, 8343
PRIOR FILING DATE: 2000-09-27	PRIOR APPLICATION NUMBER: 60/234, 274
PRIOR FILING DATE: 2000-09-21	PRIOR APPLICATION NUMBER: 60/223, 2223
PRIOR FILING DATE: 2000-09-21	PRIOR APPLICATION NUMBER: 60/228, 924
PRIOR FILING DATE: 2000-08-30	PRIOR APPLICATION NUMBER: 60/224, 518
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/236, 366
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/224, 519
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/220, 966
PRIOR FILING DATE: 2000-07-26	PRIOR APPLICATION NUMBER: 60/241, 805
PRIOR FILING DATE: 2000-10-20	PRIOR APPLICATION NUMBER: 60/249, 299
PRIOR FILING DATE: 2000-11-17	PRIOR APPLICATION NUMBER: 60/236, 327
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/241, 785
PRIOR FILING DATE: 2000-10-20	PRIOR APPLICATION NUMBER: 60/244, 611
PRIOR FILING DATE: 2000-11-01	PRIOR APPLICATION NUMBER: 60/225, 266
PRIOR FILING DATE: 2000-08-14	PRIOR APPLICATION NUMBER: 60/226, 266
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/236, 366
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/251, 855
PRIOR FILING DATE: 2000-12-08	PRIOR APPLICATION NUMBER: 60/251, 866
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/229, 344
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/234, 999
PRIOR FILING DATE: 2000-09-25	PRIOR APPLICATION NUMBER: 60/229, 343
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/229, 343
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/229, 343
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/229, 287
PRIOR FILING DATE: 2000-09-01	PRIOR APPLICATION NUMBER: 60/229, 513
PRIOR FILING DATE: 2000-09-05	PRIOR APPLICATION NUMBER: 60/231, 411
PRIOR FILING DATE: 2000-09-08	PRIOR APPLICATION NUMBER: 60/229, 505
PRIOR FILING DATE: 2000-09-05	PRIOR APPLICATION NUMBER: 60/236, 366
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/237, 033
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/237, 033
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/237, 033
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/236, 373
PRIOR FILING DATE: 2000-09-29	PRIOR APPLICATION NUMBER: 60/236, 807
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/237, 033
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/237, 033
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/237, 033
PRIOR FILING DATE: 2000-10-02	PRIOR APPLICATION NUMBER: 60/237, 033

